BTO RESEARCH REPORT 780

A review of existing methods to collect data on seabird flight height distributions and their use in offshore wind farm impact assessments

Feather, A.P., Burton, N.H.K., Johnston, D.T. & Boersch-Supan, P.H.

Birds Science People

ACKNOWLEDGEMENTS: This work was commissioned by Natural England as part of the ReSCUE (Reducing Seabird Collisions Using Evidence) project, with funding from The Crown Estate and Natural England's Species Recovery Programme. The ReSCUE project forms part of the Offshore Wind Evidence and Change (OWEC) programme, led by The Crown Estate in partnership with the Department for Energy Security and Net Zero and Department for Environment, Food & Rural Affairs. The Offshore Wind Evidence and Change programme is an ambitious strategic research and data-led programme. Its aim is to facilitate the sustainable and coordinated expansion of offshore wind to help meet the UK's commitments to low carbon energy transition whilst supporting clean, healthy, productive, and biologically diverse seas. The ReSCUE project is led by Natural England, and we are particularly grateful for the input from Andrew Harwood as the project technical lead, Alex Banks as the senior responsible officer, and Eddie Cole as the project manager. We also thank members of the ReSCUE Project Advisory Group, including the Royal Society for the Protection of Birds, Joint Nature Conservation Committee, The Crown Estate, Defra, Offshore Wind Industry Council and Marine Directorate, and attendees of an Expert Panel workshop in May 2024, who have provided valuable input and comments on the report. The contents of this report reflect the author's views, and The Crown Estate is not responsible for any use that may be made of the information it contains.

A review of existing methods to collect data on seabird flight height distributions and their use in offshore wind farm impact assessments

Report of work carried out by the British Trust for Ornithology on behalf of Natural England and the Offshore Wind Evidence and Change Programme

Feather, A.P., Burton, N.H.K., Johnston, D.T. & Boersch-Supan, P.H.

BTO Research Report 780

BTO, The Nunnery, Thetford, Norfolk IP24 2PU
Tel: +44 (0)1842 750050 Email: info@bto.org
Registered Charity Number 216652 (England & Wales), SC039193 (Scotland).

ISBN 978-1-912642-77-9

Contents

GLOSSARY

Accuracy
Accuracy refers to a measure of the difference between a measurement or estimate and a recognised standard value (e.g. the true value of a quantity). It is a measure of systematic error or bias.

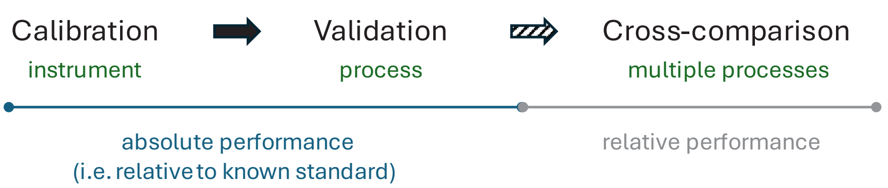
Air-gap
Air-gap refers to the space between the instantaneous sea surface and the lowest point in the circumference of the turbine blades.

Barometric altimeter
Barometric altimetry uses a barometer that is supplied with a nonlinear calibration to indicate altitude as determined based on the measurement of atmospheric pressure. The greater the altitude, the lower the pressure.

Bias
Bias refers to a systematic deviation of measurements and statistics from the truth due to characteristics of the experimental design (e.g. sampling, measurement, analysis). Hence, the extent to which a sampling, measurement or analytical method does not represent the population, value or statistic thought to be described. Here, bias is generally specific to each sampling method and can be quantified, but with varying degrees of difficulty. In practice, bias is generally measured as the difference between a raw measurement or the expected value of an estimator, and the true value of the quantity of interest.

Calibration
Calibration is an absolute assessment of the accuracy and precision of measurement values delivered by a sensor or instrument of a calibration standard with a known value (see Box 1).

Box 1: Schematic illustration of the relationship between different types or stages of monitoring technology assessments.



Collision risk model (CRM)
Collision Risk Models are used to predict the number of bird collisions that might be caused by a wind farm development. The current industry standard in the UK is a stochastic version of the Band Model, which requires input parameters describing species-specific information on biometrics, flight characteristics and the expected amount of flight activity; and turbine-specific information on blade size, blade pitch, rotor rotation period and the anticipated proportion of time that turbines will be operational.

Cross-comparison
A cross-comparison assesses the outputs of two or more technologies or processes relative to each other, regardless of whether these processes are calibrated or validated. It provides relative measures of agreement which may or may not correspond to the agreement of either method with the truth (see Box 1).

Detection probability
Detection probability quantifies a sampling method's ability to detect all available individuals.

Digital aerial survey (DAS)
Digital aerial surveys collect successive photographs or video from an aircraft as it transects the survey area.

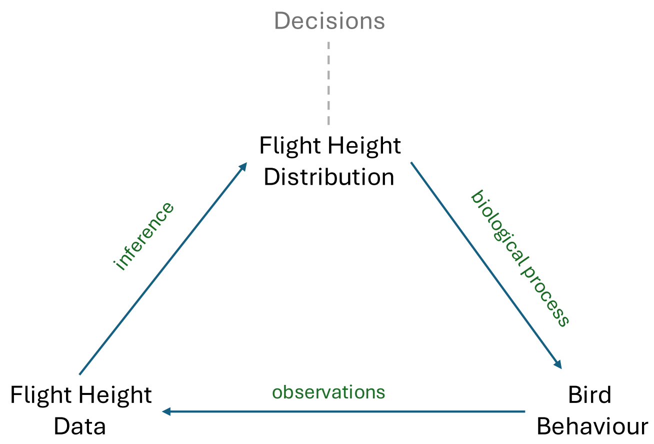
Ellipsoid
An ellipsoid is a geometrically perfect but simplistic model of mean sea level around the globe.

Eulerian
Eulerian survey designs sample at predetermined stations or along continuous transects and are often replicated through time. The primary objective of Eulerian sampling approaches is to obtain information about animal distribution and abundance in a predefined space and time.

Flight altitude, Flight height
Measures of the vertical location of a bird relative to a reference level. Flight height is generally used to describe the distance between a bird and the land or sea surface. Flight altitude is generally used to describe the distance between a bird and a specified reference value (e.g. mean sea level, chart datum). The terms are used interchangeably in much of the literature, so it is generally advised to qualify either term with a relevant reference value (see also Box 2).

Flight height distribution (FHD)
A flight height distribution is an idealised description of the bird behaviour that is realised as the frequency with which birds occupy airspace along the vertical axis. Inferences about the flight height distribution are based on flight height data, which consists of observations of said behaviour (see Box 2).

Box 2: *Bird behaviour* is a realisation of an unobservable and generally incompletely understood biological process governed by an idealised *flight height distribution*. Said behaviour can be observed – generally in an imperfect manner owing to sampling and measurement error – and the resulting *flight height data* can in turn be used to make inferences about the nature if the underling *flight height distribution* which then can be used for decision making.



Geoid
A geoid is an irregular model of mean sea level around the globe, assuming only the influence of the local gravitational field and the rotation of the Earth (i.e. no effect of landmass, wind, or tide).

Ground sampling distance (GSD)
The ground sampling distance is the distance between two consecutive pixel centres measured on the ground, i.e. the image resolution in units of the imaged surface.

Highest astronomical tide (HAT)
The highest astronomical tide is the highest sea level that can be expected to occur under average meteorological conditions and under any combination of astronomical conditions over one lunar nodal cycle (18.6 years).

Inertial measurement unit (IMU)
An inertial measurement unit is a device that can measure and report specific gravity and angular rate of an object to which it is attached. It is used to determine and record changes in pitch, roll, and yaw of a survey platform such as an aircraft or surface vessel.

Lagrangian
Lagrangian survey designs track individual animals through space and time using data logging or tracking devices.

Rangefinder
Rangefinders refer to methods that require observers to visually identify and track individual birds using optical instruments, while flight height is formally estimated using sensor-based (e.g. laser, compass, GPS, inclinometer) measurements (e.g. elevation angle, distance, bearing/azimuth) and basic mathematical principles (e.g. trigonometry). When paired with a compass and clock, laser rangefinders are sometimes referred to as ornithodolite.

Light detection and ranging (LiDAR)
LiDAR is a remote sensing method that uses light in the form of a pulsed laser to measure distances between objects and the sensor.

Lowest astronomical tide (LAT)
Lowest astronomical tide is the lowest sea level that can be expected to occur under average meteorological conditions and under any combination of astronomical conditions over one lunar nodal cycle (18.6 years).

Mean high water neaps (MHWN)
The height of mean high water neaps is the average throughout the year of two successive high waters when the tidal range is at its lowest (neap range).

Mean high water springs (MHWS)
The height of mean high water springs is the average throughout the year of two successive high waters when the tidal range is at its highest (spring range).

Mean low water neaps (MLWN)
The height of mean low water neaps is the average throughout the year of two successive low waters when the tidal range is at its lowest (neap range).

Mean low water springs (MLWS)
The height of mean low water springs is the average throughout the year of two successive low waters when the tidal range is at its highest (spring range).

Mean sea level (MSL)
Mean sea level is the datum for measurement of elevation and altitude.

Measurement error
Measurement error in this report generally refers to the precision (i.e. how close results are to one another) and accuracy (i.e. how close results are to the true value) of flight height measurements and/or related quantities associated with each method. See also sampling error.

Measurement error model
Measurement error models are statistical models that explicitly account for measurement errors in the quantities of interest.

Precision
A measure of the likely spread of repeated measurements or estimates of a quantity, hence a measure of random error.

Radio Detection And Ranging (radar)
Radar is a radiolocation system that uses radio waves to determine the distance (ranging), angle (azimuth), and radial velocity of objects relative to the site.

Rotor swept zone (RSZ)
The rotor swept zone refers to the circular area defined by the blades as the turn.

Sampling error
Statistical error that occurs when methods (e.g. experimental design, sampling method) do not select a sample that represents the entire population of data. The difference between the sample result and the population characteristic being estimated. In practice, the sampling error can rarely be determined because the population characteristic is not usually known.

Sampling volume
Sampling volume refers to the volume of airspace that is effectively sampled by a method. The shape or geometry of this volume may be complex.

Sea surface height (SSH)
Sea surface height (SSH) is the height of the sea surface above a reference ellipsoid.

Trilateration
Trilateration is the use of distances for determining the unknown position coordinates of a point of interest, often around Earth (geopositioning).

Truncation
The process of limiting consideration or analysis to data that meet specific criteria.

Validation
Validation is the assessment that a process meets its predetermined specifications and quality attributes. In the offshore monitoring context this can be 'internal' validation or 'ecological' validation which tend to be achieved in this order during the study process: we start with internal/controlled validation of monitoring approaches before validating in real ecological circumstances (see Box 1).

Executive summary

1. Significant growth within the offshore renewable energy industry (e.g. wind farms) is expected in the coming years and environmentally friendly development is a priority. Accurately predicting, mitigating, and compensating the ecological impacts of offshore development is therefore crucial to ensuring development can continue at pace without significant negative effects. Seabirds are a key component of marine ecosystems, and many species are potentially vulnerable to marine development. Offshore wind farms might impact seabird populations in several ways (e.g. displacement, barrier, and indirect effects) but direct mortality caused by collisions with turbines is of particular concern for several key species (e.g. Gannet, Kittiwake, and large gulls).

2. For the purposes of environmental impact assessment, collision rates with offshore wind farm turbine rotors are predicted using collision risk models (CRMs) which provide a quantitative estimate of risk. Information on species' flight height distributions is fundamental to predictions and continuous flight height distributions provide more robust estimates of mortality risk, compared to wider discrete flight height bands. Flight height data can be collected using a variety of sampling methods (e.g. visual surveys, photogrammetric digital aerial surveys, animal-borne tracking devices, radar, LiDAR), with which there are uncertainties (e.g. sampling error, measurement error) and logistical constraints (e.g. environmental conditions, species-specific behaviour).

3. This document presents a review of existing methods for collecting seabird flight height data and their potential to produce flight height distributions that might be used in CRMs. The strengths, weaknesses, and limitations of different methods are identified and sources of measurement and sampling error, uncertainty and bias assessed. Best practice recommendations are provided for prominent methods and how data might be best utilised to inform stakeholders is considered.

4. None of the methods reviewed could provide species-specific flight height distributions that were fully representative of the populations of interest under all relevant environmental conditions (i.e. biotic, abiotic) and across all ecologically important temporal scales (e.g. daily, seasonal, annual) of variation. Aside from animal-borne technologies, all sampling methods are vulnerable to systematic over- or underestimation of flight height in particular height bands (e.g. close to the sea-surface) which consequently impacts estimates of collision risk. Bias (positive and negative) is introduced via observer error (e.g. visual surveys), technical challenges (e.g. rangefinders), equipment limitations (e.g. radar, LiDAR, animal-borne tracking devices) and failure to account for variable detection probability within the surveyed area (all methods).

5. Measurement errors are generally better understood than sampling errors. Accuracy of individual flight height estimates from several methods (e.g. rangefinders, stereophotogrammetry, high frequency (< 20s sampling interval) animal-borne tracking devices, LiDAR, microphone array) is generally within 10 m of true flight heights in favourable conditions. Measurement errors (precision) however can be more than 100 m (e.g. visual surveys, photogrammetric digital aerial surveys, low frequency animal-borne tracking device) due to equipment characteristics (e.g. sensor accuracy, sampling frequency), human behaviour (e.g. height estimates) and from interactions between height measurements and supplementary data (e.g. sea level pressure, natural body size variation, and other reference values). Increasingly complex sampling methods (e.g. animal-borne GPS, aerial imagery) were found to simultaneously incorporate multiple sources of error which can interact to alter (e.g. inflate variance, introduce bias, distort shape) flight height distributions with consequences for CRMs.

6. The lack of robust analytical procedures for determining heterogeneity in each method's detection probabilities prevents the effectively sampled volume from being calculated for most if not all available methods. The true frequencies with which flight heights are distributed is therefore rarely estimated. Developing procedures to determine each method's detection probability is therefore a priority, particularly for methods for which the theoretical surveyed volume can be determined relatively easily (e.g. LiDAR, aerial imagery).

7. Sampling in a temporally non-biased manner (e.g. with respect to weather conditions, daylight hours) was noted as a particularly widespread challenge and was therefore also highlighted as a research priority. Telemetry studies that focus on quantifying species-specific relationships between temporally varying conditions (e.g. weather, time of day, behavioural state) and flight height are best placed to improve understanding, but novel sampling designs and the incorporation of remotely sensed data will likely be required. However, no one method provides information that is representative of all environmental conditions or of spatial variation, for a given species; thus, the integration of information across multiple measurement methods is likely to be required to provide more representative flight height distributions.

8. The continued development and assessment of methods for estimating seabird flight height distributions has significantly improved current understanding (e.g. limitations, uncertainty, collision risk). The potential accuracy of flight height estimates appears to be sufficiently high (< 10 m) to allow inferences at the vertical scales of interest (air gap, RSZ) and advanced statistical techniques (e.g. state-space models, nonlinear models) have allowed for a more rigorous quantification of uncertainty by describing the underlying distributions and providing confidence estimates. However, some methods (e.g. low-frequency GPS, aerial photogrammetry) exhibit large flight height measurement errors (> 50 m) under some, or all observation conditions and it is therefore crucial that measurement uncertainty is considered routinely. There remains a lack of agreement in the flight height estimates produced by different methods and the way many observational studies are designed is a key driver of uncertainty. Most methods were not originally designed to sample flight height distributions (e.g. radar, LiDAR), many datasets were not originally collected to describe flight height distributions (e.g. animal-borne tracking devices) and environmental limitations (e.g. rangefinders) regularly require last minute changes to experimental designs which reduce their effectiveness. There is consequently a pressing need for the development of best-practice guidelines to help ensure studies are designed robustly and data collection/reporting is standardised. Technological advances are generating a wide range of novel opportunities for flight height studies. Continued progress will require clear documentation of all practical steps (e.g. methods, analysis) and data (i.e. raw) involved to be freely available to all stakeholders.

9. Rangefinders, LiDAR, and animal-borne tracking devices (high frequency GPS) provide species-specific flight height distributions that are accurate and precise such that the underlying distributions can be statistically modelled. They are also capable of sampling prior to wind farm construction (i.e. baseline data collection) and may be scaled to regional/national operations. We therefore suggest that field validation of these methods is a useful research priority for the ReSCUE project. Other methods however can add value to current understanding (e.g. by being able to sample in inclement weather or at night) and such methods should be used where appropriate.

1. INTRODUCTION

1.1. Background

Marine ecosystems are experiencing unprecedented rates of environmental change (e.g. biodiversity, temperature, pH) due to the direct (e.g. exploitation, development, disturbance) and indirect (e.g. climate change) impact of human activities (Bugnot *et al*. 2021, Gissi *et al*. 2021, Halpern *et al*. 2008). Understanding the consequences of such change to biodiversity is fundamental to sustainable ecosystem management and essential for environmental impact assessment (Rees *et al*. 2020). Significant growth within the offshore renewable energy industry (e.g. wind farms) is expected in the coming years to meet sustainable energy targets and sustainable development is a priority (GWEC 2023). Accurately predicting, mitigating and compensating the ecological impacts of offshore development is therefore crucial to ensuring development can continue at pace without significant negative effects (Fox *et al*. 2006, Rahman *et al*. 2022, Shields *et al*. 2009).

Seabirds are a key component of marine ecosystems due to the wide range of ecological functions and ecosystem services (i.e. supporting, regulatory, cultural) they perform and provide (Grant *et al*. 2022, Mosbech *et al*. 2018, Signa *et al*. 2021). Many species are highly sensitive to environmental change (e.g. development, disturbance, climate change) and are therefore potentially vulnerable to marine development (Dias *et al*. 2019, Mitchell *et al*. 2020). Offshore wind farms might impact seabird populations through temporary (i.e. disturbance) or permanent (i.e. habitat destruction/degradation) displacement, barrier effects (e.g. migration, foraging), indirect effects (e.g. though impacts on productivity and prey resources), and collision mortality (e.g. structure, rotors). Direct mortality, caused by birds colliding with turbines, is of particular concern for some species due to the high survival and longevity of seabirds (Bailey *et al*. 2014).

Collecting direct observations of birds colliding with turbines is difficult in most wind energy settings due to the spatial and temporal extent of the effort required. Collision events are also relatively rare and indirect monitoring via carcass searches is generally preferred in terrestrial settings. There are however several well recognised sampling biases associated with carcass detection probability (e.g. carcass persistence, searcher detection rate) which must be addressed (Aschwanden *et al*. 2018, Domínguez del Valle *et al*. 2020, Huso *et al*. 2016). Considerable efforts (e.g. dummy carcasses, search dogs) have improved the accuracy with which undetected carcass numbers are estimated within onshore wind farms but the approaches used are unsuitable for highly dynamic marine environments which can quickly transport carcasses away from collision sites. Direct observations of bird collisions with turbines have been recorded in the marine environment but a large amount of effort is required and the results are likely to be site- / turbine- / season- / weather- / species-specific (e.g. Skov *et al*. 2018). For the purposes of environmental impact assessment, potential collision rates within offshore wind farms are predicted using collision risk models (CRMs) which provide a quantitative estimate of risk (Masden & Cook 2016). The standard CRM used within the UK is a mechanistic model based on the probability of a turbine blade occupying the same space as a bird flying through the turbine rotor swept volume (Band, 2012a, 2012b, 2000, Band *et al*. 2007). The probability of collision is predicted using the physical features (e.g. wingspan, body length), flight characteristics (e.g. speed, height) and behavioural traits (e.g. nocturnal activity, avoidance) of the bird, and the blade dimensions (e.g. width, length, pitch), structural properties (e.g. rotor speed, hub height) and operating schedule (e.g. activity) of the turbine.

Information on species' flight height distributions is fundamental to estimating the collision risk of seabirds with offshore wind farms as part of the environmental impact assessment process (Largey *et al*. 2021, Masden *et al*. 2021). The original CRM (here after 'basic Band model') assumes a uniform distribution of flight heights across the rotor swept zone (RSZ) and requires single estimate of the proportion of birds at risk height (Band 2000, Band *et al*. 2007). In this case, flight height data can be collected and summarised at relatively coarse resolutions as the minimum information required is whether an individual is flying above or below two specific thresholds (i.e. the lower and upper limits of the RSZ, e.g. Lane *et al*. 2020). Collision risk can also be predicted using an extended version of the model (here after 'extended Band model') which expects continuous flight height distributions that quantify the relative frequency at 1 m height bands between 0 m and 500 m (Band 2012a, 2012b). The extended Band model generates more refined estimates of collision mortality by accounting for variation in bird density with altitude (and therefore probability of collision) across the risk area but requires more detailed data collection from which vertical density distributions of flying birds can be produced (Johnston *et al*. 2014, 2023). A deterministic version of the Band model utilises point estimates of flight height distributions, whilst the more recent stochastic version incorporates variation based on bootstrapped sampling of each species' flight height distribution (Caneco *et al*. 2022, McGregor *et al*. 2018).

Robust estimates of collision risk require data from sampling methods that accurately describe the complete range of heights that a species can occupy and the frequency with which it does so. Flight height data can be collected using a variety of sampling methods (e.g. visual surveys, digital aerial surveys, animal-borne tracking devices, radar, LiDAR) with which a number of uncertainties (e.g. sampling error, measurement error) and logistical constraints (e.g. environmental conditions, species-specific) are associated (Largey *et al*. 2021, Searle *et al*. 2023, Thaxter *et al*. 2015). Seabird flight characteristics (e.g. location, speed, direction, height) additionally vary in response to environmental conditions (e.g. wind speed, temperature), animal behaviour (e.g. foraging, commuting, resting), individual-based traits (e.g. species, age, sex, breeding status) and in response to human activities (e.g. development, fishing) which creates considerable variation (spatial, temporal) that must be captured within the sampling process (Ainley *et al*. 2015, Lane *et al*. 2020, van Erp *et al*. 2023). There is consequently substantial ambiguity over the reliability of estimated flight height distributions and predicted collision rates.

This review forms part of the ReSCUE (Reducing Seabird Collisions Using Evidence) project. This overarching project aims to provide confidence in flight height data, its use in impact assessments, and development of effective mitigation solutions by:

1. Reviewing and collating reliable sources of complementary seabird flight height data;
2. Providing confidence in survey methods and commissioning additional surveys to address knowledge gaps;
3. Promoting access to data and facilitating the collection of new data to agreed standards;
4. Providing user-friendly tools and guidance for the interrogation and application of the collated data to facilitate speedy, reliable, impact assessments;
5. Examining what factors influence seabird flight height, vulnerability to collision, and requirements for mitigation;
6. Updating evidence for cumulative impact assessments based on improved evidence; and
7. Developing mitigation principles and guidelines to reduce impacts on vulnerable species and improve the consenting process.

This review thus forms the first element of Objective (1) above, to build on existing reviews to appraise methods in relation to the collection of seabird flight height data. Subsequent work will propose best practice in relation to the collection and analysis of these data for use in impact assessments.

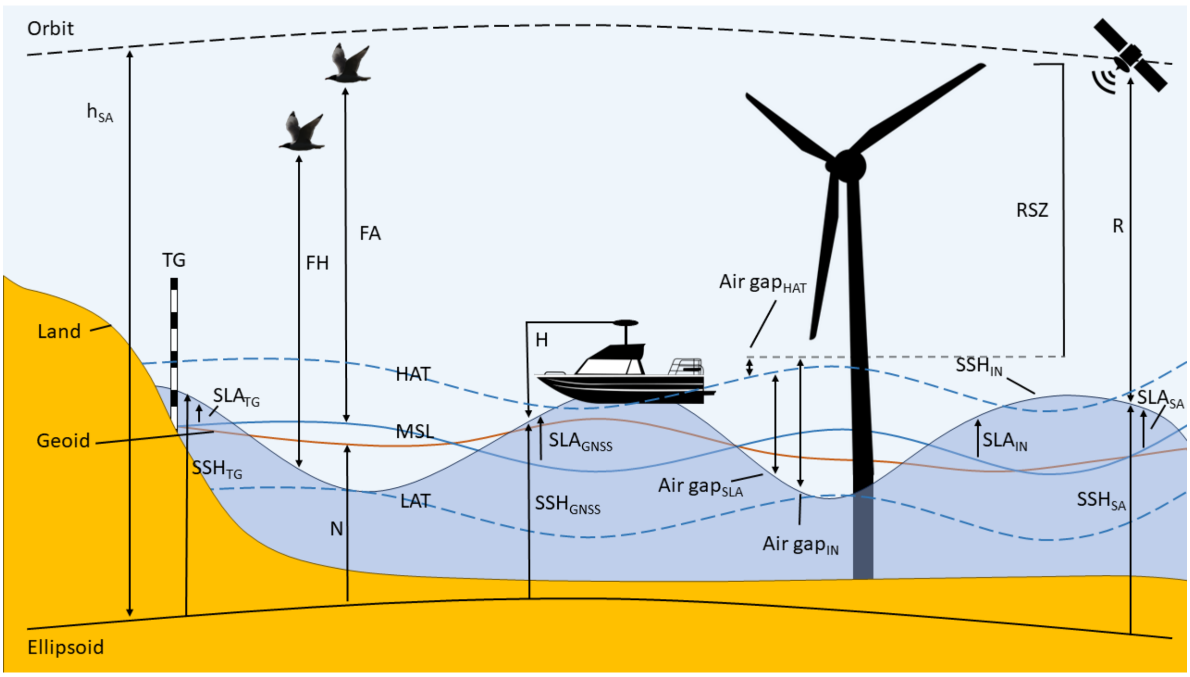
1.2. Conceptual foundations of flight height measurements and models

1.2.1. Vertical frame of reference

Measures of seabird flight height in the marine environment refer to the distance between a bird in flight and some measure of the sea surface. However, the sea surface height as experienced by marine birds relative to structures (e.g. wind turbines) which are generally fixed to the seabed, varies temporally (e.g. by day, month, year) and spatially (e.g. inshore, offshore) and experiences periodic extremes (i.e. tide). Sea surface height (SSH) is typically defined using standard terms to summarise the variation over one lunar nodal cycle (18.6 years, Figure 1.1, Liibusk *et al*. 2020, Peng *et al*. 2019, Tamisiea *et al*. 2014). Mean sea level (MSL) is the average height of the sea surface over one cycle or in the absence of tides. Mean high water (MHW) and mean low water (MLW) are the average of all the daily high and low water levels observed over one cycle. Mean high water spring (MHWS) and mean low water spring (MLWS) are the average maximum and minimum heights when the tidal range is greatest (i.e. spring range). Mean high water neap (MHWN) and mean low water neap (MLWN) are correspondingly the average maximum and minimum heights when the tidal range is lowest (i.e.

neap range). The highest (HAT) and lowest astronomical tides (LAT) are the maximum and minimum heights that are predicted to occur under average meteorological conditions over one cycle. Some sampling methods (e.g. visual surveys, rangefinders, barometric altimetry) estimate flight height directly in relation to the sea surface (Harwood *et al.* 2018, Johnston *et al.* 2023) while others (e.g. GPS, aerial imagery) estimate flight altitude relative to constant reference values (e.g. MSL, Cook *et al.* 2018, Harwood *et al.* 2018, Johnston & Cook, 2016, Johnston *et al.* 2023). Some methods (e.g. LiDAR) can estimate either flight height or flight altitude (Cook *et al.* 2018, Wicikowski *et al.* 2022). From a biological perspective flight height measurement should be (as close as possible) related to the instantaneous sea level, the resulting values can then be related to HAT.

Figure 1.1. Schematic plot of the relationship between measures of sea surface height (instantaneous – SSHIN, satellite altimetry – SSHSA, tide gauge – SSHTG, GNSS – SSHGNSS), two commonly used reference altitudes (mean sea level – MSL, a modelled geoid) and a wind turbine. The instantaneous sea level anomaly (i.e. difference between measured and reference value) is shown for each method (instantaneous – SLAIN, satellite altimetry – SLASA, tide gauge – SLATG, GNSS – SLAGNSS). Measures of flight height (FH) refer to the distance between a flying bird and the instantaneous sea surface height while measures of flight altitude (FA) refer to the distance between a flying bird and a reference altitude. Turbine characteristics (e.g. RSZ, Air gap) are measured from the highest astronomical tide (HAT), the rotor swept zone (RSZ) refers to the area between the upper and lower limits of a turbines blades and the air gap refers to the distance between the RSZ and the HAT (Air gapHAT). Both the instantaneous air gap (Air gapIN) and the difference between this and the HAT value (Air gapSLA) are also shown. Additional abbreviations include: hsa - ellipsoidal height, N – geoid undulation, R – distance between the satellite and the sea surface, H – GNSS antenna height from the sea surface. At the tide gauge, the geoid undulation (N) approximately coincidences with the vertical datum zero N ≈ mean sea surface height above ellipsoid.



Flight altitudes are typically estimated relative to a geometrically perfect (i.e. ellipsoid) but simplistic model of MSL (e.g. World Geodetic System 1984, WGS84). They may alternatively be estimated relative to an irregular (i.e. geoid) model of MSL if it was only influenced by the local gravitational field and the rotation of the Earth (i.e. no effect of landmass, wind, or tide, e.g. OSGM15). Flight altitudes are also sometimes estimated relative to empirical measures of sea level as measured from a reference point over a reference period (e.g. Ordnance Datum Newlyn, defined as the MSL as recorded by the Newlyn Tidal Observatory between 1915 and 1921). Flight heights are converted to altitude using simultaneously observed or predicted values of SSH. There are many methods by which SSH (e.g. satellite altimetry, tide gauges, buoys) is routinely measured and data are available at a variety of spatial and temporal scales (Liibusk *et al.* 2020, Tamisiea *et al.* 2014).

The height characteristics of wind turbines (e.g. RSZ, air-gap) are usually measured relative to the HAT (a precautionary approach that is relevant for the engineering specs of the infrastructure), but the structures are usually fixed to the seabed and the perceived height of structures by seabirds varies relative to SSH (Figure 1.1). The height of the collision risk area (i.e. RSZ) relative to the sea surface is therefore continually decreasing and increasing as the tide rises and falls, or relative to local wave/swell height. Depending on the extent of SSH variation in an area of interest, collision risk heights may therefore also need to be corrected using simultaneously observed or predicted values of SSH. This is currently achieved using a site-specific tidal offset between HAT and MSL.

1.2.2. Sources of error and uncertainty

Flight height data should ideally be species-specific and representative of all individual traits (e.g. age, sex, body size), behaviours (e.g. foraging, commuting, resting), annual cycle stages (i.e. breeding, non-breeding, migration) and environmental conditions (e.g. temperature, windspeed, precipitation, food) that can be encountered within the area of interest. Measurements should be free from bias, accurate and precise such that flight height distributions characterise the complete range of heights and the true frequency with which they occur. The types of error and uncertainty associated with seabird flight height estimates can be broadly grouped into sampling and measurement processes, respectively. Sampling error is generally systematic and emerges both in response to the experimental design (e.g. non-random sampling) and the method of data collection (e.g. sampling geometry, detection accuracy, detection probability). Measurement error refers to the precision of the height measurements obtained from detected birds (i.e. how repeatable the measurements of the same target are) and accuracy (i.e. how close the obtained measurement is to the true value) associated with each method.

Data sampling methods (and associated error/uncertainty) can be grouped into two categories (Eulerian and Lagrangian) based on experimental design (Largey *et al*. 2021, Phillips *et al*. 2019, Watanuki *et al*. 2016). Eulerian sampling methods (e.g. visual surveys, aerial imagery, LiDAR) collect observations in a predetermined spatial and temporal frame of reference (e.g. coordinates, transects). Lagrangian experimental design collects data at discrete spatial and temporal locations using animal-borne tracking devices. The primary difference between Eulerian and Lagrangian sampling is the level of inference each method can achieve about populations or individuals, respectively, and/or spatiotemporal domains of interest.

1.2.2.1. Sampling error

There is considerable uncertainty associated with each method's ability to adequately sample the population of interest and none of the methods discussed can provide species-specific flight height distributions that are fully representative of the populations (i.e. properties, constituents), environmental conditions (i.e. biotic, abiotic) and temporal scale (e.g. decades) they aim to describe.

All Eulerian observation methods fundamentally sample a finite volume of airspace (hereafter 'sampled volume'), and the maximum sampled volume is typically determined by sensor characteristics. Examples of the theoretically sampled volume for different methods and the associated variation in sampling error/efficiency is provided (Figure 1.2). The sampled volume of aircraft-based surveys (i.e. aerial imagery, LiDAR) increases with increasing distance from the sensor due to the triangular vertical cross section of the surveyed volume, which is governed by the aperture angles of the employed cameras or LiDAR devices (Cook *et al*. 2018, Johnston & Cook 2016, Figure 1.2a). The sampled volume of radar-based approaches is constrained by the shape of the radar beam, the horizontal cross section of which generally increases with distance from the equipment (Schmid *et al*. 2019, Figures 1.2d and e). The sampled volume associated with human observer-based methods (e.g. visual surveys, rangefinders) is often assumed to be uniform within certain distance limits but is generally not well understood (Figure 1.2c).

Many Eulerian sampling methods are biased (positively and negatively) towards low altitudes due to challenges associated with observations close to the sea surface. During visual surveys, observers are just as likely to assign birds to the incorrect height band as to the correct band and routinely underestimate (i.e. positive observation bias) flight height (Harwood *et al*. 2018, Perrow *et al*. 2017). Rangefinders improve the accuracy at which observers estimate flight heights, but difficulties associated with targeting low (< 10 m) flying individuals can result in relatively few records (i.e. negative observation bias; Borkenhagen *et al*. 2018). Both radar and LiDAR (when not paired with cameras) are prone to increased false-positive detection rates due to reflections from waves or spray and consequently data below a threshold height (e.g. 2 m) have historically been routinely removed (Cook *et al*. 2018, van Erp *et al*. 2024). The result is a systematic over- or underestimation of flight height in certain altitude bands which consequently affects estimates of collision risk.

In addition to complex geometries of sampling volumes, the detection probability of each Eulerian method is generally not uniform across the respective sampled volumes and effective sampling rates are therefore variable. The detection probability of visual-based methods (e.g. visual survey, rangefinder, photogrammetry) decreases with increasing distance from the observer or sensor (i.e. negative observation bias) due to limitations associated with visibility, optical resolution and/or targeting individuals (Barbraud & Thiebot, 2009, Borkenhagen *et al*. 2018, Harwood *et al*. 2018). The detection probability of radar and LiDAR-based approaches decreases with increasing distance (i.e. negative observation bias) because the energy of return signals becomes too low to detect (Dokter *et al*. 2013, May *et al*. 2017). The detection probability of sound-based approaches (e.g. microphone array) also decreases with increasing distance from the sensor as sounds become too low to detect (Stepanian *et al*. 2016). The result is a systematic over- or under-estimation of flight height and collision risk.

Both the sampled volume and detection probability of each Eulerian method may also vary considerably (temporal and spatial) in response to environmental conditions. Sampling during poor weather and/or low light conditions remains particularly challenging due to impacts on the measurement process and/or deployment limitations of equipment (Largey *et al*. 2021). The sampled volume and detection probability of visual and imagery-based methods decreases with worsening weather conditions and interference from the sea surface and/or precipitation can similarly alter the effective sampled and detection probability volume of radar- and LiDAR-based systems. Detection probabilities are also fundamentally related to the size of birds, larger individuals can be detected at greater distances relative to smaller ones (Barbraud & Thiebot, 2009, Cook *et al*. 2018, Schmid *et al*. 2019).

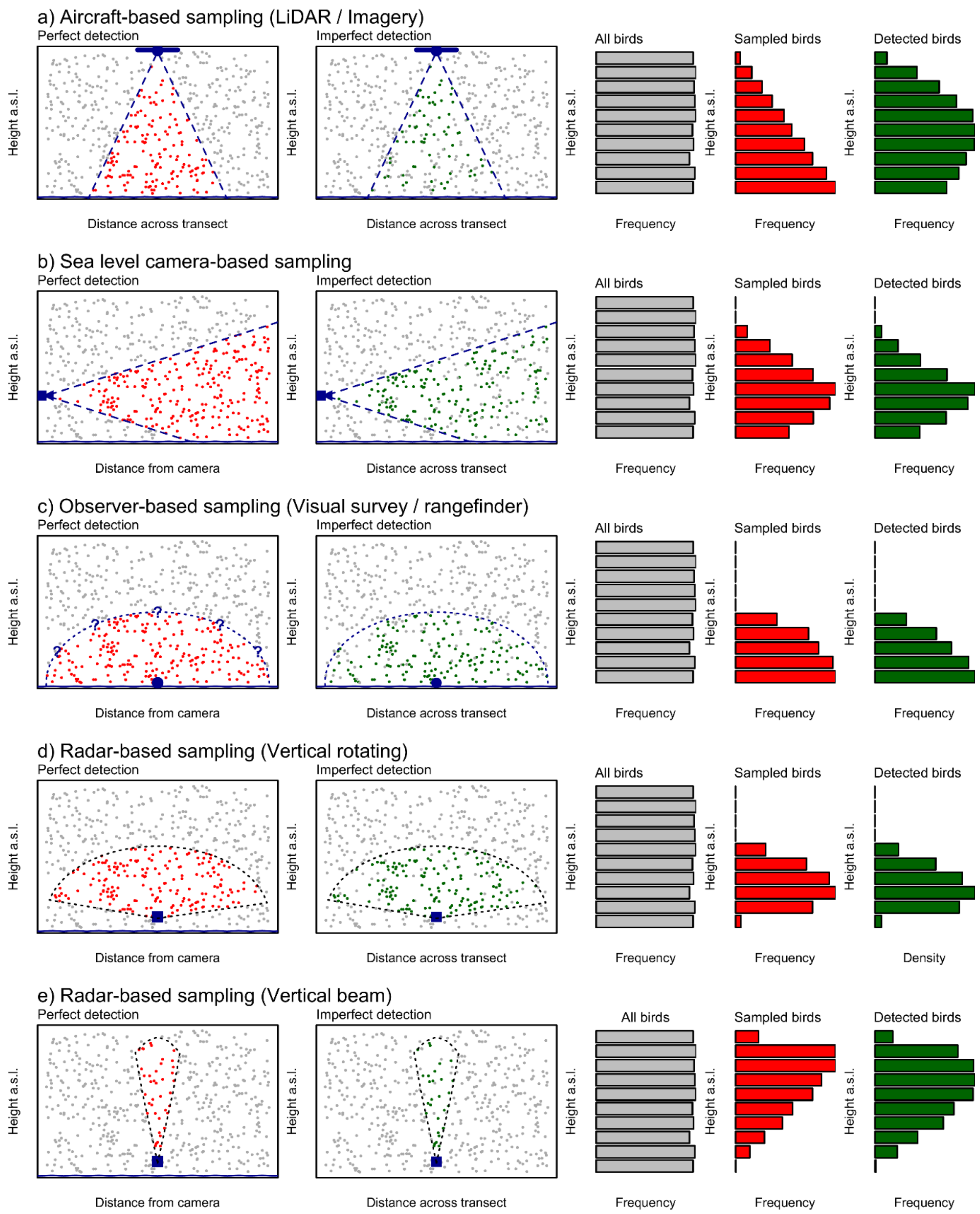
Where digital data is collected in relatively large volumes (e.g. radar, imagery, LiDAR, animal-borne tracking device) many of the processing steps (e.g. detection, classification, tracking) are increasingly automated (e.g. image-processing algorithms, target-tracking algorithms, classification algorithms). While the algorithms must intermittently miss or reject true observations and accept false observations, the frequency with which it occurs is largely unknown (Urmy & Warren 2020). The detection probability of the various algorithms is also likely to vary in space (e.g. habitat, altitude) and time (e.g. weather, time of day), particularly in marine environments (e.g. waves, sea spray). The same uncertainties are also present where processes are not automated. There is for example a reliance on humans for the interpretation of images (e.g. target detection, species identification, matching imaged individuals with LiDAR points) and the accuracy with which such processes occur is not well understood.

1.2.2.2. Measurement error

There is also considerable uncertainty surrounding the accuracy and precision with which each method measures flight height. All methods incorporate some degree of vertical error, often from multiple sources (e.g. equipment, operator, supplementary data) and at various stages (e.g. data collection, data analysis) of the sampling process. As the complexity of sampling methods increases, errors can arise and interact in increasingly complex and counterintuitive ways which further complicates all subsequent inferences.

Measurement errors that arise due to equipment characteristics (e.g. sensor accuracy, sampling frequency) are generally inherent to data collection and usually generate random noise around height estimates (Harwood *et al*. 2018, Lato *et al*. 2022). Those that result from human behaviour (e.g. height estimates) are also introduced during data collection and can introduce both random noise and systematic bias to height estimates (Harwood *et al*. 2018, Perrow *et al*. 2017). Errors that result from interactions between height measurements and supplementary data (e.g. sea level pressure, reference values) are introduced while data are processed and may introduce both random noise and systematic bias to height estimates (Boersch-Supan *et al*. 2024, Johnston *et al*. 2023, Schaub *et al*. 2023). Increasingly complex sampling methods (e.g. animal-borne GPS, aerial imagery) can simultaneously incorporate multiple sources of error (e.g. device, supplementary data) which interact to alter (e.g. inflate variance, introduce bias, distort shape) flight height distributions (Boersch-Supan *et al*. 2024, Péron *et al*. 2020, Ross-Smith *et al*. 2016).

Figure 1.2 Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for a) aerial survey, b) forward looking camera, c) human observer, d) vertically rotating radar and e) vertical radar beam. Scatterplots depict the sampled volume (cross section) which varies in relation to the sensor/observer and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



1.2.3. Flight height distribution models and data integration

The uncertainties in the sampling and measurement processes discussed above interact to generate samples of flight height observations that do not faithfully represent the true distribution of flight heights. The latter therefore have to be reconstructed by means of statistical modelling. Models for flight height distributions have been proposed for about a decade, initially to synthesise flight height observation across multiple studies (Johnston *et al*. 2014) and subsequently to explicitly account for observation errors (Johnston & Cook, 2016, Ross-Smith *et al*. 2016, Peron *et al*. 2017, Fleming *et al*. 2020, Davies *et al*. 2024). The latter is achieved using so-called state-space modelling frameworks which conceptually separate the biological process (i.e. the selection of flight heights by the bird) from the observation process (i.e. the sampling and measurement characteristics of the employed observation technology). This generally requires distributional assumptions about both the shape of the true flight height distribution (e.g. a log-Normal distribution (Ross-Smith *et al*. 2016) or Gamma mixture (Johnston & Cook 2016)), and the nature of the sampling and measurement errors (e.g. Gaussian errors (Ross-Smith *et al*. 2016) or Student t errors (Peron *et al*. 2017)). While state-space models in principle allow the joint estimation of parameters for both the process and observation models, there are fundamental identifiability constraints (Auger-Methe *et al*. 2016), which can seriously limit the applicability of these models to real-world data. This is in particular the case when measurement errors are of a similar or larger magnitude as the variance of the flight height distribution – which is the case for several observation techniques such as low-frequency sampling GPS devices or aerial photogrammetry (e.g. Lato *et al*. 2022, Boersch-Supan *et al*. 2024). Error calibration of the observation technology and independent empirical validation of measurement error modelling approaches is therefore crucial to assessing whether or not robust biological inferences can be drawn for any particular dataset (Fleming *et al*. 2020, Boersch-Supan *et al*. 2024).

Published state-space models for flight heights generally only consider a single observation process, i.e. a single observation technology. However, the state-space framework naturally expands beyond this to multiple observation processes, as the modular model structure in principle allows the specification of an observation model for each data input. Such modelling approaches, so-called model-based data integration, are rapidly gaining traction in other realms of ecological statistics to combine different data sources in a single statistical model (e.g. animal count, detection/non-detection, and presence-only data; Pacifici *et al*. 2017, Isaac *et al*. 2020, Mancini *et al*. 2022). We anticipate to adapt this approach to draw inferences about flight height distributions from multiple datasets in the course of the ReSCUE project.

1.3. Aims and objectives

The overall aim of this document is to review existing methods for collecting seabird flight height data and their potential to produce flight height distributions that might be used in CRMs. The review aims to:

- Identify strengths, weaknesses, and limitations of different observation methods.
- Assess sources of measurement and sampling error, uncertainty, and bias for each method.
- Consider how data are best utilised to inform scientists, regulators, and stakeholders.

2. METHODS

2.1. Literature review

The methods used to collect data on seabird flight heights have been previously reviewed in detail (Largey *et al*. 2021, Thaxter *et al*. 2015). For the purposes of this project the literature review was not systematic but, to ensure transparency and complete reporting, a general guide to the approach used is provided. The literature search considered sources published prior to April 2024 and was limited to sources that were accessible online and written in English. Web of Science, Google Scholar, Google Search, and the Tethys online database (www.tethys.pnnl.gov) were searched for reports, peer-reviewed publications, book chapters and theses. The primary search terms used were 'seabird', 'bird' or 'avian' in combination with 'flight height' and 'flight altitude'. Particular attention was given to previous reviews of flight height data collection methods and the references therein (Largey *et al*. 2021, Thaxter *et al*. 2015). All studies were investigated for species-specific flight height estimates and information concerning the strengths, weaknesses, limitations, and sources of uncertainty associated with each sampling method was extracted.

The results are organised by methodology with a brief description of how flight heights are estimated, the characteristics of each sampling design and the error (i.e. sampling, measurement) associated with each. Key analytical and operational considerations for each method are also described. The uncertainties, strengths and weaknesses are then broadly discussed in relation to estimating collision risk with offshore turbines before recommendations for best practice is provided and key knowledge gaps and research priorities are identified (summarised in Table 4.1). Data collection methods have previously been grouped into two categories (sensor, non-sensor, Largey *et al.* 2021). The authors considered sensor data to be those collected remotely using devices (e.g. LiDAR, radar, GPS) that transmit and/or receive signals from which flight parameters can be directly measured. Non-sensor data were considered those collected locally by observers (e.g. visual surveys) or devices (e.g. camera, microphone) which require flight parameters to be calculated post hoc with a degree of observer subjectivity. In the present report, this distinction becomes increasingly unclear as sensor technology is currently employed in one form or another in almost any approach. The distinction is therefore not used, and methods are discussed independently.

3. RESULTS

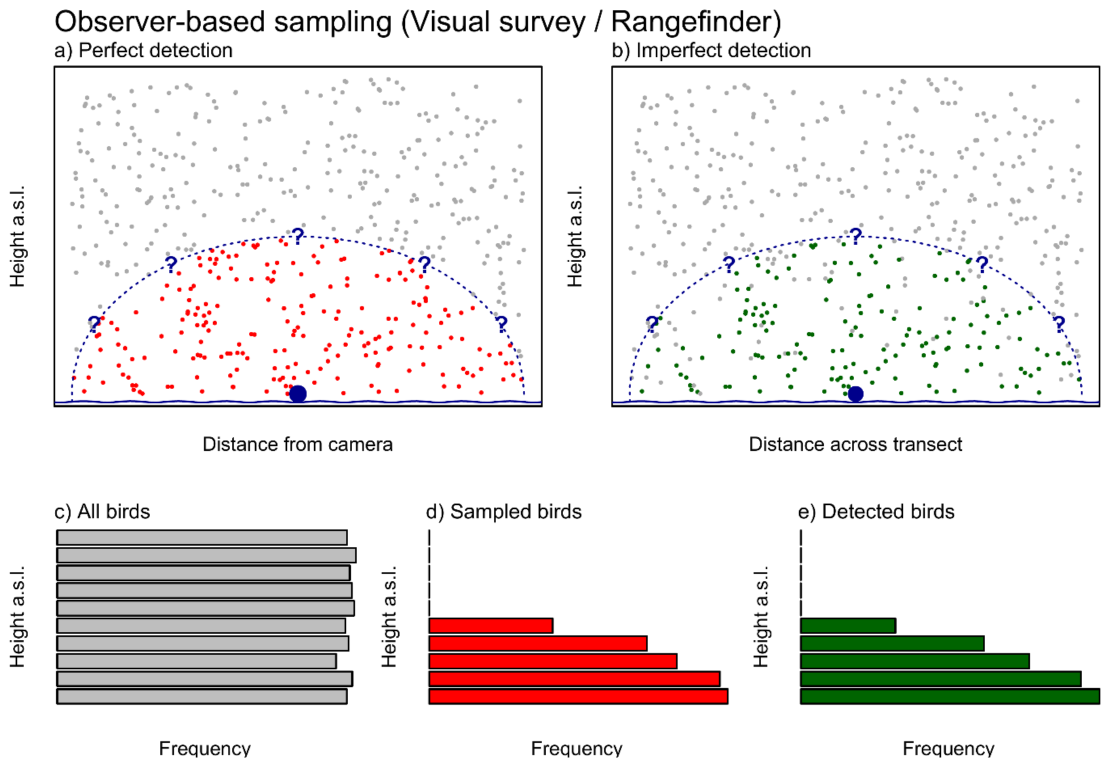
3.1. Visual surveys

Seabird flight heights have traditionally been collected via visual surveys which require trained observers to estimate the height of all birds within a set distance (e.g. < 300 m) of their location (e.g. vessel, vantage point) at predetermined (e.g. five minute) intervals (Camphuysen *et al.* 2004). Values are categorical (i.e. in height 'bands') and either set with reference to the height of fixed objects (e.g. vessel mast, turbine) or to standardised categories (e.g. Larsen & Guillemette, 2007, Leemans *et al.* 2022, McClure *et al.* 2021, Mendel *et al.* 2014, Rothery *et al.* 2009, van Bemmelen *et al.* 2022).

Sampling characteristics
Visual surveys generally follow a Eulerian experimental design in which observations are taken at predetermined locations within the area of interest (e.g. transects, vantage points, Table 4.1). Some approaches combine aspects of both Eulerian and Lagrangian methods by tracking individuals from either vessels or platforms to reconstruct movement tracks (Akeresola *et al.* 2024, Perrow *et al.* 2017, 2011). Sample sizes are generally moderate, and data can be collected for multiple species simultaneously. Visual surveys are positively biased towards mild (e.g. calm, dry, good visibility) weather conditions and daylight hours, they are non-invasive but the survey platform can induce behavioural responses (e.g. attraction, displacement) which typically introduce negative observation bias within flight height estimates (Borberg *et al.* 2005, Jarrett *et al.* 2021, Schwemmer *et al.* 2011).

The volume of air sampled via visual surveys is assumed to be a hemisphere in shape with a radius of approximately < 300 m from the observer (Figure 3.1). In practice the shape and size of the sampled volume is poorly understood. The use of trained and experienced observers results in most available individuals being uniformly detected and correctly identified to species level at horizontal distances < 100 m. Detection probability decreases with increasing distance from observers, bird size and wave height but increases with the number of observers (Barbraud & Thiebot 2009, Ronconi & Burger 2009, Spear *et al.* 2004). Detection probabilities at distances < 300 m for example, are reported to be 0.87, 0.73 and 0.69 for large (2–11 kg), medium (0.5–1.5 kg) and small (< 0.5 kg) birds respectively (Barbraud & Thiebot 2009). Additional studies demonstrate that between 20 and 80% of pursuit diving birds can be detected within a 300 m wide transect (150 m on either side of the vessel, Ronconi & Burger 2009). Methods to optimise the survey area (e.g. strip width) and analytical approaches such as distance sampling aim to correct for detection biases within two-dimensional surveys (Hyrenbach *et al.* 2007, Ronconi & Burger 2009). Equivalent concepts and methods are currently lacking for three-dimensional data collection and the effective sample volume for visual surveys is therefore generally unknown.

Figure 3.1. Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for observed-based sampling methods. Scatterplots depict the sampled volume (cross section) which is assumed to be uniform < 300 m from observers and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



Measurement characteristics

Observers typically calibrate flight height estimates with reference to the height of fixed objects (e.g. a ship's mast) but few studies have assessed the accuracy at which seabird flight heights are estimated using visual surveys. Estimates from visual surveys have however been compared with those from laser rangefinders (Harwood *et al.* 2018, Perrow *et al.* 2017).

Comparisons of estimates made using visual surveys and laser rangefinders found observers assign birds to the correct height band 30–58% of the time and differences (50–86%) among bird groups indicate species-specific drivers of measurement error (Harwood *et al.* 2018, Perrow *et al.* 2017, Thaxter *et al.* 2015). Observers assigned birds to either the same or adjacent 5 m band 92–96% of the time but routinely underestimated flight heights, particularly at greater heights (Harwood *et al.* 2018, Perrow *et al.* 2017). Visual surveys therefore likely overestimate the proportions of birds flying at low altitudes (i.e. positive observation bias) and underestimate those numbers flying at greater altitudes (i.e. negative observation bias).

Analytical considerations

Visual surveys estimate flight heights relative to the in situ sea surface and values must be converted to MSL for use in CRMs using simultaneous observations or predictions of SSH (i.e. instantaneous sea level anomaly). The effectively sampled volume is not well understood for human observers. Detection of all available individuals is generally assumed, but there is little evidence whether this assumption is fulfilled with respect to both horizontal and vertical distance (i.e. altitude) from the transect. Horizontal detection probabilities have been estimated using conventional distance sampling methods and these findings suggest that detection probabilities are near one within 100 m of the transect but decline beyond (Barbraud & Thiebot 2009, Ronconi & Burger 2009). However, the adequacy of conventional distance sampling analysis has been questioned for shipboard seabird surveys (and other stationary or slow-moving platforms that sample moving animals; Glennie *et al*. 2015, 2021). To our knowledge, there is no well-established statistical method to analyse data where the detection probability varies in both horizontal and vertical direction, although initial work on related problems (e.g. fisheries acoustics) can be found in the statistical literature (Cox *et al*. 2011, Borchers & Cox 2017).

Flight height distributions collected via visual surveys are categorical (i.e. assigned to fixed height bands) and are therefore most easily applied to the basic Band model, assuming survey height bands align with the RSZ. Depending on the vertical resolution of survey height bands, confidence intervals for flight height distributions and/or the proportion of birds at risk can be challenging to calculate, as are inferences about continuous, flight height distributions (Cook *et al*. 2012, Johnston *et al*. 2014). The combination of survey data from different studies can help in this case, particularly when height band categories vary across studies but may be less informative about underlying continuous distributions where height bands are uniform throughout (Johnston *et al*. 2014).

Operational considerations
Visual surveys are conducted both from moving platforms (typically surface vessels) and stationary vantage points, and can be deployed at various spatial (e.g. local, regional, national) and temporal (e.g. days, weeks, months) scales. The protocols for estimating seabird flight heights using visual surveys are well established and can therefore be implemented in relatively short times (Camphuysen *et al*. 2004, Tasker *et al*. 1984). Visual surveys can take place pre-, intra-, and post-construction of wind farms, they have no long-term maintenance requirements but are restricted to fine weather and daylight hours. Species-specific data can be collected for multiple species simultaneously and additional behavioural (e.g. foraging, commuting), environmental (i.e. biotic, abiotic) but limited individual-based (e.g. sex) data can be simultaneously collected. Their cost incorporates the requirements of observers (e.g. training, safety, living), their equipment (e.g. optics, Personal Protective Equipment — PPE) and the survey vehicle (e.g. fuel, crew), which can be considerable.

3.2. Rangefinders
Rangefinders refer to methods that require observers to visually identify and track individual birds using optical instruments, while flight height is formally estimated using sensor-based (e.g. laser, compass, GPS, inclinometer) measurements (e.g. elevation angle, distance, bearing/azimuth) and basic mathematical principles (e.g. trigonometry). Laser rangefinders typically estimate the slope distance and degree of inclination (elevation angle) to specified targets (e.g. birds) using a laser beam (e.g. Borkenhagen *et al*. 2018, Fijn & Collier 2022, Harwood *et al*. 2018, Leemans *et al*. 2022). The estimated values are then used to calculate the horizontal and vertical distance, and if paired with a compass or high-resolution GPS, a three-dimensional (i.e. x, y, z) position or track (e.g. Harwood *et al*. 2018, Borkenhagen *et al*. 2018, Perrow *et al*. 2017). When paired with a compass and clock, laser rangefinders are sometimes referred to as ornithodolites (Pennycuick 1982, Pennycuick *et al*. 2013). Ornithodolites incorporate a measure of azimuth and can provide additional estimates of flight speed (Cole *et al*. 2019, Largey 2020). There are many different types of rangefinders and the capability (e.g. accuracy, precision, range) and cost of each can vary considerably.

Sampling characteristics
Rangefinders generally follow a Eulerian experimental design in which observations are taken at predetermined locations within the area of interest (e.g. transects, vantage points, Table 4.1). As with visual surveys, animals can be tracked (i.e. Eulerian/Lagrangian) to reconstruct movement tracks (e.g. Cole *et al*. 2019, Perrow *et al*. 2017). Rangefinders can theoretically measure flight height at distances of up to 12 km but, as with visual surveys, observers may only reliably detect and correctly identify individuals to species level at distances < 100 m and detection probability decreases with distance and bird size (Barbraud & Thiebot 2009). Sample sizes are generally moderate (100s to 1000s), and data can be collected for multiple species simultaneously (Borkenhagen *et al*. 2018, Harwood *et al*. 2018, Skov *et al*. 2018). Laser rangefinder surveys are positively biased towards mild (e.g. calm, dry, good visibility) weather conditions and daylight hours, they are

non-invasive but the survey platform can induce behavioural responses (e.g. attraction, displacement) which introduce observation bias within flight height estimates (Jarrett *et al*. 2021, Schwemmer *et al*. 2011).

As with visual surveys, the volume of air sampled via rangefinders is assumed to be a hemisphere in shape, but the size is poorly understood (Figure 3.1). The detection process for rangefinders is composed of two steps, the observer must first visually detect a target before locking the measurement mechanism onto it. Visual detection is not considered a problem within 300 m of an observer but measuring the height of an unmanned aerial vehicle (UAV) (dimensions – 88.7 × 88.0 × 37.8 cm) was reported to be difficult at distances > 100 m and altitudes > 75 m using a rangefinder due to difficulties locking onto the target (Harwood *et al*. 2018). The probability of successfully locking onto targets is positively related to the size of individual birds and varies depending on the angle of view, the orientation (e.g. broadside) and behaviour (e.g. flapping, gliding) of the target individual, and additional environmental factors (e.g. background composition, atmospheric moisture content, physical structures). Flight height data collected via rangefinders are therefore likely to be negatively bias towards more distant or smaller individuals (Borkenhagen *et al*. 2018, Cole *et al*. 2019, Harwood *et al*. 2018, Kahlert *et al*. 2012).

Measurement characteristics

Estimates from laser rangefinders have been compared with those from UAVs, barometric altimetry, GPS triangulation and known reference points (Borkenhagen *et al*. 2018, Harwood *et al*. 2018, Largey, 2020, Prinsloo *et al*. 2021, Skov *et al*. 2018).

The accuracy and precision of a laser rangefinder (Forestry Pro, Nikon, Tokyo, Japan) was evaluated by comparing estimates with those from a M200 quadcopter (DJI, Shenzhen, China, Harwood *et al*. 2018). Overall accuracy (mean error ± standard deviation) was < 1 m (n = 407, -0.4 m ± 1.3) and improved (-0.04 m ± 0.7) when comparisons were limited to a maximum distance and altitude of 100 m and 50 m, respectively. Accuracy was found to vary nonlinearly in relation to height and distance suggesting additional factors contribute to variability in rangefinder accuracy. Precision (judged by the modelled confidence intervals) declined with increasing height but varied with distance.

The accuracy of an ornithodolite (based on Vectronix Vector 21 Aero) was assessed by comparing estimates with those from stationary objects (building and turbine) and a UAV (DJI Phantom 4) piloted at variable heights (10–120 m) and distances (50–300 m, Largey 2020). Accuracy decreased with increasing distance during stationary (distance = 50 m, mean error range = 0–2 m; distance = 5,000 m, mean error range = -2–6 m) tests. The accuracy of the ornithodolite also decreased with increasing height during mobile tests but did so at greater rates for smaller distances.

The precision of an ornithodolite (Vectronix USMC Vector 21) was assessed by examining the standard deviation of estimates from a stationary object measured at variable distances (50 m to 5 km, Cole *et al*. 2019). The study also examined how the maximum measurable distance varied in relation to body size by measuring the distance to in flight birds (n = 4200, species = 151). The precision of measurements increased with distance and was approximately 1—2 m at distances < 2 km. The maximum measurable distance increased with bird body mass.

Flight heights of Lesser Black-backed Gulls were estimated using a laser rangefinder (Vector 21 Aero) and compared with those recorded using animal-borne GPS devices (Borkenhagen *et al*. 2018). The flight height distribution obtained via the rangefinder (n = 1,785, max = 431 m, median = 21 m, min = -2 m) was less variable compared to the animal-borne GPS devices (n = 705, max = 735 m, median = 8 m, min = -10 m) and resulted in a greater proportion of values in the collision risk zone (rangefinder: 70.0% < 30 m, 29.6% 30 – 150 m and 0.4% > 150 m; GPS devices: 59.3% < 30 m, 17.0% 30–150 m and 5.7% > 150 m).

Rangefinders (Vectronix 21 Aero) were calibrated by measuring the distance to known reference points during a bird collision and avoidance study, accuracy was considered ± 10 m (Skov *et al*. 2018).

Analytical considerations

Rangefinders estimate flight height relative to the sea surface and values must be converted to MSL using simultaneous observations or predictions of SSH. The surveyed volume and detection probabilities have not been quantified. Laser rangefinders generally provide more detailed data (i.e. continuous height and distance estimates) than visual surveys and are therefore in principle better able to be used for the estimation of continuous flight height distributions, which can be directly input into both the basic and extended Band models. The underlying distribution and associated confidence intervals can also, in principle, be statistically modelled but no examples could be found during this review. However, the same caveats about decreasing detection probabilities with distance from transect and/or altitude apply. Rangefinder data are more similar to sensor data that have been modelled using 3D distance sampling in other contexts (e.g. fisheries acoustics, Cox *et al*. 2011), however, in addition to the concerns raised about the adequacy of conventional distance sampling methods for visual observer data above, the two step detection process for rangefinders adds additional analytical challenges, as target-locking occurs with a lag to the initial detection, thereby violating another key assumption of conventional distance sampling models (Glennie *et al*. 2015).

Operational considerations

Rangefinder surveys are deployed both from moving platforms (typically surface vessels) and stationary vantage points. Sampling can take place at various spatial (e.g. local, regional, national) and temporal (e.g. days, weeks, months) scales but surveys are restricted to fine weather and daylight hours. Species-specific data can be collected for multiple species simultaneously and additional behavioural (e.g. foraging, commuting), environmental (i.e. biotic, abiotic) but limited individual-based (e.g. sex) data can be simultaneously collected. Sampling can occur pre-, intra-, and post-construction of wind farms but rangefinders ideally require a stable platform of known height from which they must be calibrated. Measurements can be influenced by metal structures and atmospheric conditions. High specification devices will automatically store measurements internally (e.g. using a memory card) or can be connected (e.g. by cable, Bluetooth) to external devices for further analysis but low specification units tend to not have this capability. Handheld units need to be kept vertical to ensure the clinometer functions correctly. This is difficult in practice and could be a source of additional error.

3.3. Single-camera photogrammetry (aerial imagery)

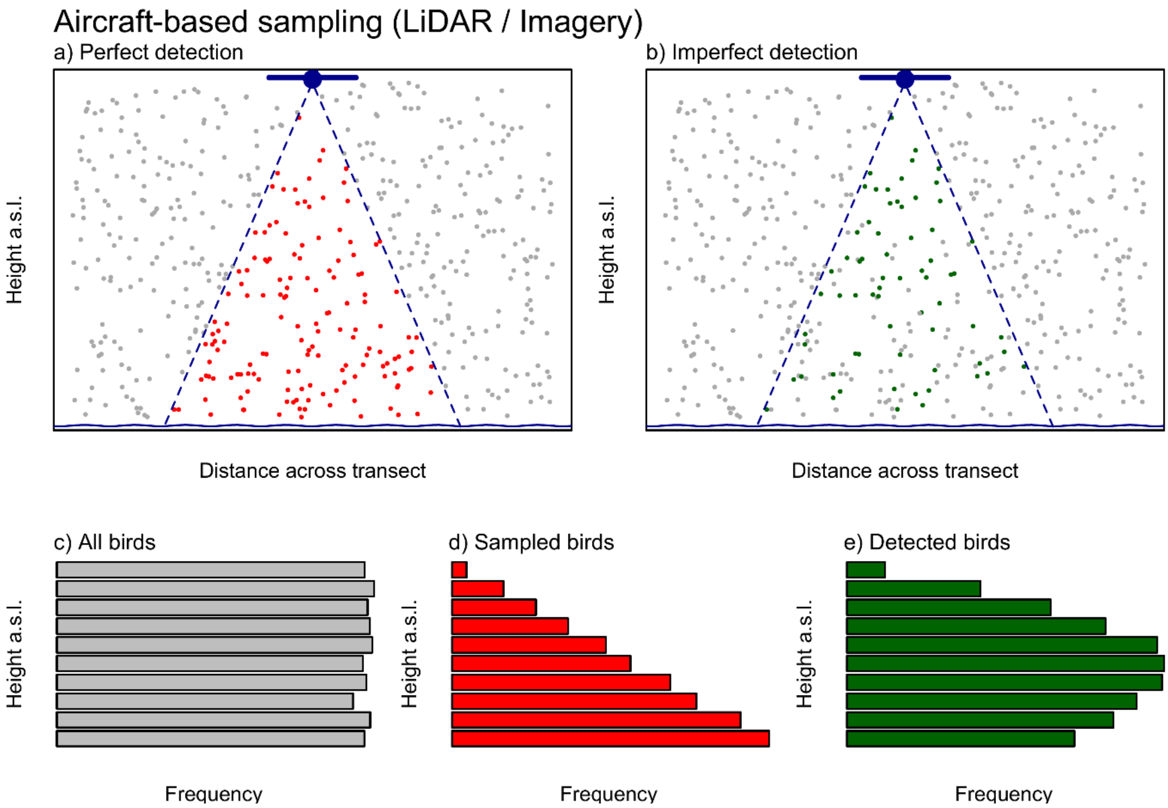
The use of digital aerial surveys (still photographs and video) to collect data on seabird flight height is relatively new (Forster *et al*. 2024, Humphries *et al*. 2023, Johnston & Cook 2016, Srinivasan *et al*. 2022). The aircraft first transects the survey area using cameras to collect data before all individual birds in the images or videos are identified to species and their flight height estimated. For still photographs or video stills, flight height estimates assume that the size of the bird is directly proportional to the distance from the camera lens. The size the imaged bird would be at sea level is first estimated using species-specific reference sizes and the ground sampling distance (GSD; i.e. pixel resolution in ground distance units) of the camera. The GSD varies with distance from the camera and flight height is estimated by scaling the height of the plane (as estimated via the vehicles Inertial Measurement Unit, IMU) using the ratio between the mean reference bird size at sea level and size in the image. This approach is highly sensitive to natural body size variation in seabirds, and it remains to be demonstrated that it is in fact capable of achieving satisfactory levels of accuracy and precision given the large intraspecific body size variation in many seabird taxa (Boersch-Supan *et al*. 2024). Alternative approaches for video footage have estimated the flight height of individual birds by comparing the speed at which the bird passes the plane to the speed of the sea surface. This is calculated for each successive pair of video frames that contain an individual bird and the mean height across each pair is used as the estimate (Cook *et al*. 2016). However, methodological details of this approach, and hence measurement characteristics, are unclear.

Sampling characteristics

Digital aerial surveys follow a Eulerian sampling design whereby the cameras are fitted to aircraft which follow predetermined transects (constant ground speed) within the area of interest (Table 4.1). Aerial surveys are non-invasive, and aircraft are piloted at altitudes (e.g. 350–600 m) are thought to minimise disturbance (Thaxter & Burton 2009). Threshold disturbance values however originate from abundance estimates (i.e. counts) and are based on whether birds at the sea surface are prompted into flight. Relatively small movements (e.g. 10 m) in response to aircraft may considerably alter flight height distributions but no information could be found regarding the three-dimensional response of birds in flight to low flying aircraft. Values also represent the maximum distance that birds can be from the aircraft (i.e. sea surface) but individuals in flight will be closer. Aircraft are fast moving observation platforms, allowing for near-synoptic coverage of large areas while generating moderate sample sizes (100s to 1,000s). Flights are biased towards clement (e.g. calm, dry, good visibility) weather conditions and daylight hours (unless infra-red capable which limits species identification).

The sampled volume of aerial imagery increases with distance from the aircraft due to the pyramidal view from the lens and birds at high altitudes are therefore less likely to be observed (i.e. negative observation bias, Figure 1.2). Observed flight height distributions therefore must be reweighted according to the geometry of the sampled volume but it is unclear whether this is routinely done. Variation in the aircrafts flight altitude should also be incorporated when determining sample volume but the extent to which altitude varies is not clear (Certain & Bretagnolle 2008). Disturbance effects are also likely to decrease with increasing distance from the aircraft but fine-scale three-dimensional responses of birds in flight to aircraft is poorly understood. Aerial surveys are thought to achieve near complete detection of medium to large-bodied seabirds and can identify most individuals to species level (Buckland *et al*. 2012). However, up to 35% of imaged birds were discarded prior to analysis in a recent study because their posture precluded accurate body size determination (Humphries *et al*. 2023), and it remains to be demonstrated whether this removal of detected birds is random with respect to flight height (Boersch-Supan *et al*. 2024). Paired LiDAR/DAS surveys would be suitable to assess whether this is a significant issue.

Figure 3.2 Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for aircraft-based sampling methods. Scatterplots depict the sampled volume (cross section) which increases with distance from the sensor and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



Measurement characteristics

The accuracy and precision of flight heights obtained via digital aerial surveys have not been fully validated. Humphries *et al*. (2023) present validation against uncertainty caused by aircraft altitude variation and/or pixelation. However, they failed to recognise that the large natural intra-specific variation in seabird body size (e.g. sex) creates a fundamental challenge when estimating flight heights using single-camera aerial imagery (Boersch-Supan *et al*. 2024, Humphries *et al*. 2023). Every imaged bird that is smaller than the reference individual will be assigned a negatively biased flight height as it appears to be further away from the camera. Likewise, for larger birds there will be a positive bias in the estimated height, and hence estimated flight height distributions based on mean reference body sizes are much more dispersed than the true underlying distributions. In practice, the uncertainty about individual flight heights caused by natural body size variation is much larger than any other error source (i.e. uncertainty about apparent body size because of pixelation, or uncertainty about aircraft height and/or attitude). This error has not been quantified experimentally, but the theoretical impact of body size uncertainty has been assessed via simulations (Boersch-Supan *et al*. 2024). Measurement error increased with body size variation (expressed as the coefficient of variation, CV) to values > ± 50 m when species identity was known and > ± 200 m when birds were classified as unknown gulls. Forster *et al*. (2024) present a revision of the approach proposed by Humphries *et al*. (2023). The revised methodology provides reduced errors in survey-level mean flight heights but is not able to deliver individual-level flight height estimates, therefore severely reducing the ability to draw direct inferences about collision risk. Instead of using individual-level flight height estimates they propose an indirect modelling approach which uses survey-level mean flight heights combined with strong assumptions about the underlying flight height distribution to estimate collision risk. This revised approach is yet to be independently validated, as for potential sampling errors created by the exclusion of birds in certain postures, paired LiDAR/DAS surveys may be able to shed further light on the feasibility and robustness of this analysis approach.

Analytical considerations

Aerial imagery estimates flight height relative to the altitude of the aircraft which is estimated via GPS. Flight heights are therefore estimated in relation to a reference MSL and may need conversion before use in CRMs. The theoretical surveyed volume can be estimated using camera specifications and aircraft position (e.g. longitude, latitude, altitude). The detection probability of aerial imagery is currently not quantified but thought to be near perfect for many seabird taxa (but not all, detection probabilities are likely to be smaller than 1 for small and/or dark coloured species such as European Storm Petrels (*Hydrobates pelagicus*) or small shearwaters (Baker *et al*. 2022, Certain & Bretagnolle 2008). Aerial imagery generates continuous flight height distributions which can be input directly into both the basic and extended Band models. Confidence intervals can be calculated manually for still images by examining potential uncertainty in the size of imaged birds or known variation in reference values. Confidence intervals are calculated for video images by bootstrapping different pairs of frames and calculating a new mean for each sample. The underlying distribution can, in principle, be statistically modelled to account for measurement error, although it is questionable whether existing measurement error models (e.g. Johnston & Cook 2016) are robust in the face of the magnitude of measurement errors associated with this observation approach.

Operational considerations

Digital aerial surveys are well suited to cover large spatial areas near-synoptically and can be used pre-, intra-, and post-construction of wind farms or far offshore where the use of other methods may not be feasible. They are however restricted to higher altitudes where aircraft operate over wind farms due to safety regulations. The subsequent change to sampling characteristics (e.g. ground sampling distance, identification rates, sample volume, detection probability) needs to be carefully considered for any pre–post construction comparison survey work. Surveys are restricted to fine weather and daylight hours; the data are species-specific and can be collected for multiple species simultaneously. Additional, but limited, behavioural data (e.g. flight speed) are embedded in the images or videos, but no environmental data or individual characteristics can be simultaneously collected.

3.4. Stereophotogrammetry

Stereophotogrammetry is used to estimate three-dimensional coordinates of points within photographic images (still or video). It involves two or more cameras setup at known positions (e.g. altitude, latitude, longitude) so that multiple overlapping images can be recorded simultaneously. The directional angle (horizontal and vertical) of each camera lens is used to determine the three-dimensional location (i.e. triangulation) of points within the images. The technique was initially developed for topographical mapping but is now regularly used to measure and analyse three-dimensional structure within many ecosystem types (e.g. riparian, forests, dunes, biogenic reefs) and at a variety of spatial scales (Pulido Mantas *et al.* 2023). Stereophotogrammetry and/or imagery-based ornithodolites have been successfully used to create three-dimensional tracks of animal movement and there are multiple commercial systems and software (e.g. IdentiFlight, Bioseco, spoor.ai, ORJIP, 2022) that are currently in operation or development within terrestrial and offshore wind farms in relation to bird collisions (de Margerie *et al.* 2015, Evangelista *et al.* 2017, Ling *et al.* 2018). The method is therefore in principle suitable for collecting flight heights of birds (Prinsloo *et al.* 2021).

Sampling characteristics

Stereophotogrammetry uses a Eulerian experimental design whereby areas of interest can be continuously (i.e. 24 h d^{-1}) monitored from predetermined locations (Table 4.1). Flight heights are species-specific and depending on the instrumentation used, the approach may be restricted to clement (e.g. calm, dry, good visibility) weather conditions and daylight hours (unless infra-red capable). Equipment can be manually, or motion activated, and sample sizes will vary depending on the survey location and duration. Stereophotogrammetry is not invasive but the structures on which cameras are mounted may induce behavioural responses (e.g. attract) and create negative bias within flight height estimates.

Both the sampled volume and detection probability of stereophotogrammetry depends on the cameras (e.g. focal length, angle of view), the system setup (e.g. distance between cameras, pointing angle of cameras) and atmospheric conditions (Figure 3.3). Few data are available describing sampling characteristics, but birds can in principle be detected at distances of several hundred meters or more (e.g. Duerr *et al.* 2023, Linder *et al.* 2022, Rolek *et al.* 2022).

Measurement characteristics

Although, under ideal conditions, stereophotogrammetry can achieve 3D track reconstructions with cm-scale accuracy and precision, even for fast moving birds of prey (Brighton *et al.* 2022, 2017, Prinsloo *et al.* 2021), performance in typical monitoring applications is more variable and strongly depends on technical aspects of both the employed cameras and analytical workflows. There are several potential sources of measurement error associated with stereophotogrammetry. Pointing errors in direction from the attitude (i.e. pitch, roll, heading) sensor contribute to position error as a function of distance with greater range to target leading to larger errors. Stereo pair-based errors result from choosing slightly different points in the images and/or lack of perfect temporal synchronisation between image pairs which leads to an incorrect disparity measurement. The precision of the estimated range is proportional to the baseline (distance between cameras) which directly affects how this error is propagated through to the final measurement. A larger baseline reduces positioning error but also increases the minimum operating distance before a target is within the stereo overlap area. Sequential flight heights of moving individual birds are additionally likely to be temporally autocorrelated (Prinsloo *et al.* 2021).

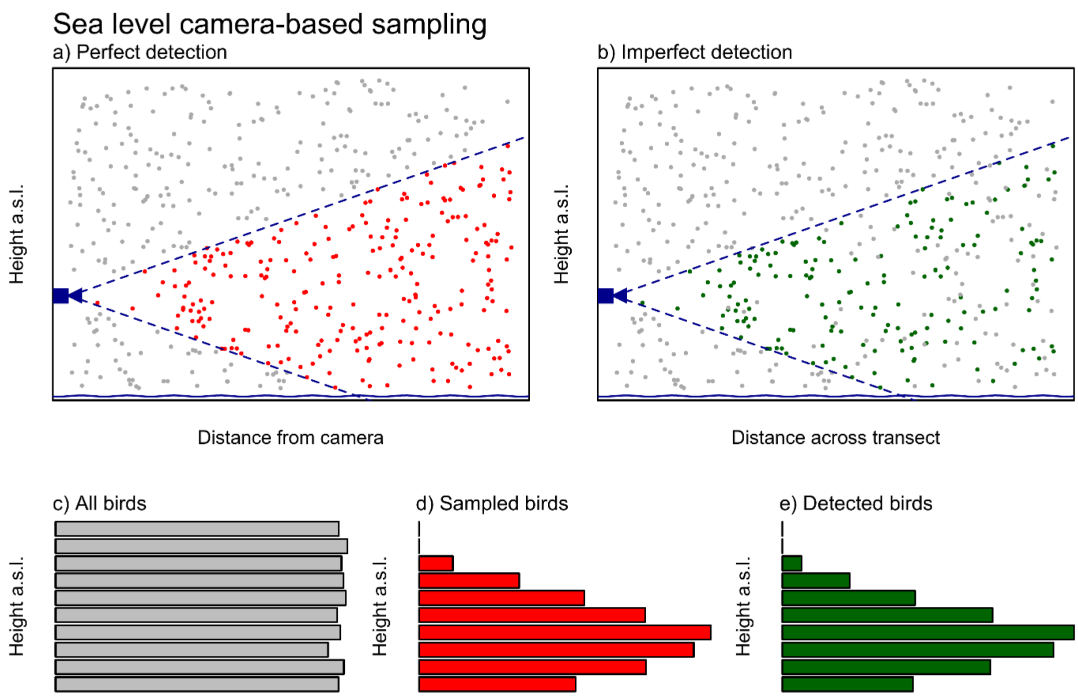
The accuracy and precision of stereo and video photogrammetry was assessed by comparing values with those from stationary locations (e.g. structures, targets), moving objects (e.g. ball), UAVs and rangefinders (Clausen *et al.* 2023, de Margerie *et al.* 2015, Prinsloo *et al.* 2021).

The accuracy (mean height m ± standard error) of stereophotogrammetry was assessed by measuring the height of known structures and comparing values with those from laser rangefinders (Prinsloo *et al.* 2021). Precision (mean standard deviation ± standard error) was assessed using the standard deviation from estimated bird flight heights. There was no statistically significant difference between stereophotogrammetry and laser rangefinders when measuring three static structures with mean estimated heights generally within < 1 m between different methods at distances up to approximately 200 m. Photogrammetrically measured flight heights (n = 316) were precise to 0.07 ± 0.05 m up to 275 m, within 1 m at 400 m and measurable up to 535 m.

The position accuracy (relative distance error, RDE) of an anti-collision system (Bioseco) was tested by comparing values (distance) with those from UAVs (1.5 and 2.0 m wingspan) and rangefinders (Clausen *et al.* 2023). A RDE of approximately 10% (UAV flight 1 = 13%, UAV flight 2 = 9%, UAV flight 3 = 11%) was determined for a maximum range of 530 m.

The three-dimensional accuracy (RMS) of stereo videography was assessed by comparing values with stationary (flags: 30, 50, 70, 90, 110 m) and moving points (tennis ball: 0, 40, 60, 80, 100, 120 m) at known distances (de Margerie *et al.* 2015). Spatial uncertainty was observed to be < 1 m within 300 m of the observer and < 0.1 m within 100 m of the observer (de Margerie *et al.* 2015).

Figure 3.3 Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for horizontal camera-based sampling methods. Scatterplots depict the sampled volume (cross section) which increases with distance from the sensor and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



Analytical considerations

Stereophotogrammetry estimates flight height relative to the position of cameras which is provided by GPS. Flight heights are therefore estimated relative to a reference MSL and may require conversion for use in CRMs. Flight height distributions collected via stereophotogrammetry are continuous and therefore derived distributions can be applied to both the basic and extended Band models. The underlying distribution and associated confidence intervals may also be statistically modelled but no examples could be found during the present review. High-precision 3D reconstructions of bird flight trajectories (e.g. Brighton *et al.* 2022, Prinsloo *et al.* 2021) typically involve manual or semi-manual image processing, this prevents scalability of this method to long-term monitoring. However increasingly automated image analysis is used in this context, but the performance of such analysis pipelines is less well understood.

Operational considerations

Stereo- or multi-camera systems are currently mostly used to provide continuous monitoring at fixed locations. Cameras require structures on which they can be mounted and are therefore restricted to intra- and post-construction, all equipment requires maintenance (hardware and software) while in operation and there are considerable post-processing requirements (i.e. image analysis). Stereophotogrammetry collects species-specific data and can do so for multiple species simultaneously. Data collection is restricted to clement weather and daylight hours. Additional behavioural data are embedded in the images or video and environmental data loggers can potentially be installed alongside cameras.

3.5. Microphone array

Microphone arrays are increasingly employed to quantify the three-dimensional position and movement of sound producing animals (Dutilleux *et al.* 2023, Rhinehart *et al.* 2020). Multiple time-synchronised microphones (i.e. array) are deployed from which the animals locations is determined by quantifying the time delay of the sounds arrival (Dutilleux *et al.* 2023). The method has been used to track the movements of animals in flight (e.g. birds, bats) and can therefore potentially be used to document the flight heights of seabirds and/or nocturnally migrating land birds at offshore installations (Dutilleux *et al.* 2023, Gayk & Mennill, 2020, Stepanian *et al.* 2016, Suryan *et al.* 2016).

Sampling characteristics

Microphone arrays monitor from fixed and predetermined locations thereby following a Eulerian sampling design (Table 4.1). They can continually (i.e. 24 h d^{-1}) monitor the surrounding airspace and identify multiple individuals to species level simultaneously. Microphone arrays are non-invasive but the structures on which microphones are mounted may induce behavioural responses (e.g. attract) and negatively bias flight heights. Sample sizes will vary depending on the location, duration, and time of year.

The sampled volume of microphone arrays is theoretically complex depending on the capabilities, direction, and number of microphones in the array. Microphone arrays will only detect birds that are actively producing sound while in flight. This limits the approach to certain taxa, and periods at which individuals are actively calling (e.g. migration). The detection probability decreases with distance from the equipment as sound becomes too low to detect. Detection distance was assessed using a kite fitted with GPS and speakers from which sounds were transmitted (Stepanian *et al.* 2016). The array could reliably detect calls $<$ 90 m above ground level before the signal extinguished into the ambient noise. Detection probability will therefore also depend on the characteristics (e.g. pitch, volume, tone) of sounds and the sampling volume will likely be species-specific. Microphones are also prone to false positive observations in the direction of background noise (e.g. insects, leaves, water, wind) but this may be minimised by mounting microphones in parabolic containers (Dutilleux *et al.* 2023, 2023, Gayk & Mennill 2020, Stepanian *et al.* 2016).

Measurement characteristics

The location of each microphone is determined via GPS devices and the associated measurement error (horizontal and vertical) will be introduced into flights height values. The accuracy and precision of microphone arrays has been assessed by comparing values to those of stationary (e.g. mounted speakers) and moving targets (e.g. kite lofted GPS with speakers (Gayk & Mennill 2020, Stepanian *et al.* 2016).

The accuracy and precision of a microphone array (six microphones fixed to three poles) was assessed by comparing values (mean difference) to those of a kite fitted with a GPS and speakers from which bird calls were transmitted (Stepanian *et al.* 2016). The maximum horizontal and vertical distances from the centre of the microphone array to the kite were 105 and 140 m, respectively. Vertical accuracy was found to be ± 5 m and ± 10 m for 60.1% and 80.4% of observations respectively, values were consistently underestimated.

The accuracy and precision of a microphone array (eight microphones fixed to four poles) was assessed by comparing values (mean difference ± standard error) to those of a stationary speaker mounted 10 m off the ground toward the centre of the array (Gayk & Mennill 2020). The speaker played test tones and flight calls for which mean location accuracy was 1.52 ± 0.34 m and 2.04 m ± 0.37 m, respectively.

Analytical considerations
The survey volume and detection probability of microphone arrays is unreported for many studies but will likely depend on the species and be influenced by environmental conditions. Distance sampling approaches are available in principle for such data (Marques *et al.* 2013, Pérez-Granados & Traba, 2021), but similar caveats about variable detectability in three dimensions, as raised in the visual observations section theoretically apply. Flight height is estimated relative to the position of microphones which is provided by GPS and/or static topographic datums. They are therefore estimated relative to a reference MSL and may require conversion for use in CRMs. Flight height distributions collected via microphone arrays are continuous and therefore can be applied to both the basic and extended Band models. The underlying distribution and associated confidence intervals may also be statistically modelled but no examples could be found during the present review.

Operational considerations
Microphone arrays are generally used to continuously monitor at fixed locations. Individuals are identified to species levels and data can be collected for multiple species simultaneously. Microphones require structures on which they can be mounted and are therefore restricted to intra- and post-construction. Devices require maintenance (hardware and software) while in operation and significant post-processing once data are collected. Additional information (e.g. sex) may be embed in sound recordings and environmental loggers can be deployed alongside each array. Data collection may be restricted to calm weather but can occur during both daylight and nocturnal hours. The present examples of three-dimensional animal tracking occurred within terrestrial environments. Microphone arrays have been trialled within the offshore environment to monitor migration (e.g. abundance, diversity) and detect collisions but no examples of measuring flight height could be found in the present review (Farnsworth & Russell 2007, Suryan *et al.* 2016).

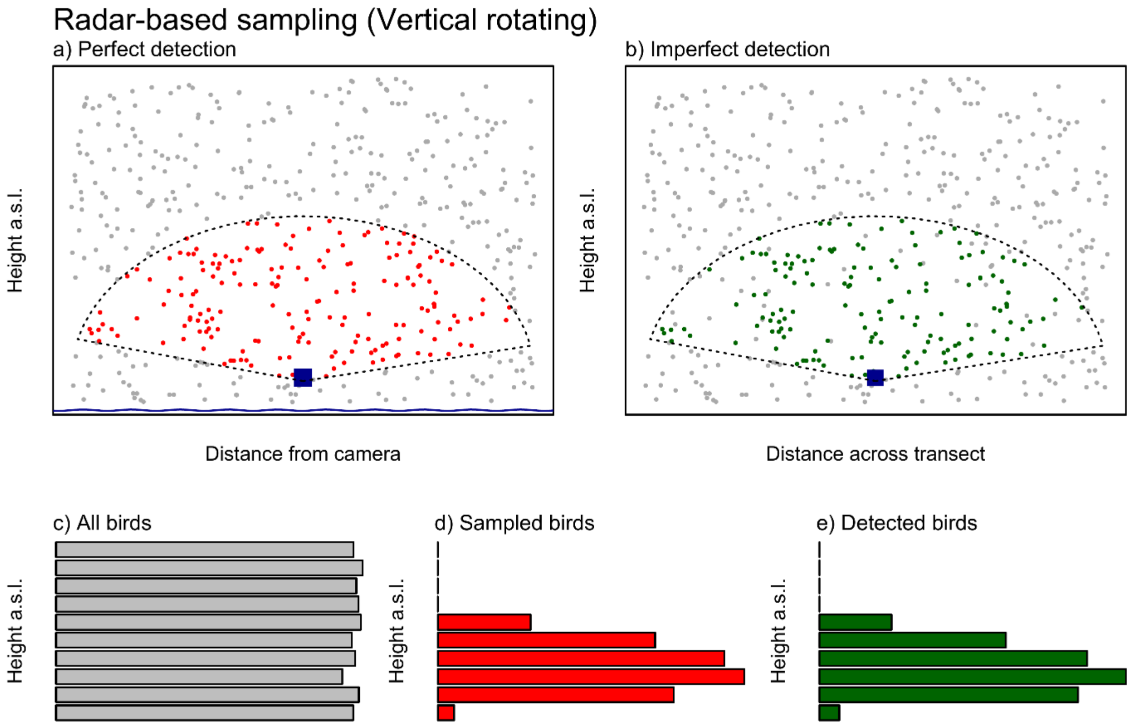
3.6. Radio Detection And Ranging (radar)

Radio detection and ranging (radar) is a radiolocation system that uses radio waves to determine the distance, angle, and radial velocity of objects relative to the site. Distances are estimated using known wave speeds and the time taken for each reflection to return. Radars can be used to map the trajectory, density, and distribution of moving objects (i.e. birds) and have been used to study the flight of birds (e.g. passerines, raptors, waders, seabirds) for many decades (Hüppop *et al.* 2019, Shamoun–Baranes *et al.* 2019). Three broad categories of radar system are regularly used to estimate the flight height distribution of birds: weather radars, marine radars and dedicated bird radars (Hüppop *et al.* 2019, Nilsson *et al.* 2018). Weather radars are usually constructed as part of a wider, nationwide network for long-term monitoring of atmospheric conditions (Cohen *et al.* 2022, Dokter *et al.* 2011, Kranstauber *et al.* 2020, Weisshaupt *et al.* 2021). Marine and dedicated bird radars are smaller, often portable systems which are typically used to monitor local site-specific conditions.

Weather radars are commonly situated on towers from which they intermittently (5–10 minutes) scan (360°) at multiple fixed elevation angles (e.g. 0.5–19.5°). They have the largest horizontal (< 250 km) and vertical (< 5 km) range but altitude distributions are typically derived for a small proportion (5–25 km) of this where the vertical resolution (e.g. 200 m) can be resolved into the altitude patterns of interest (Kranstauber *et al.* 2020). The lower altitudes associated with wind turbine collisions (< 200 m) are generally not well monitored due to ground clutter (i.e. non-bird echoes) and other challenges of interpreting echoes in the near field of the antenna. Weather radar are consequently more suited for sampling flight height at large spatial (e.g. migration flyway) and temporal (e.g. years) scales (Cohen *et al.* 2022, Dokter *et al.* 2011, Nilsson *et al.* 2018).

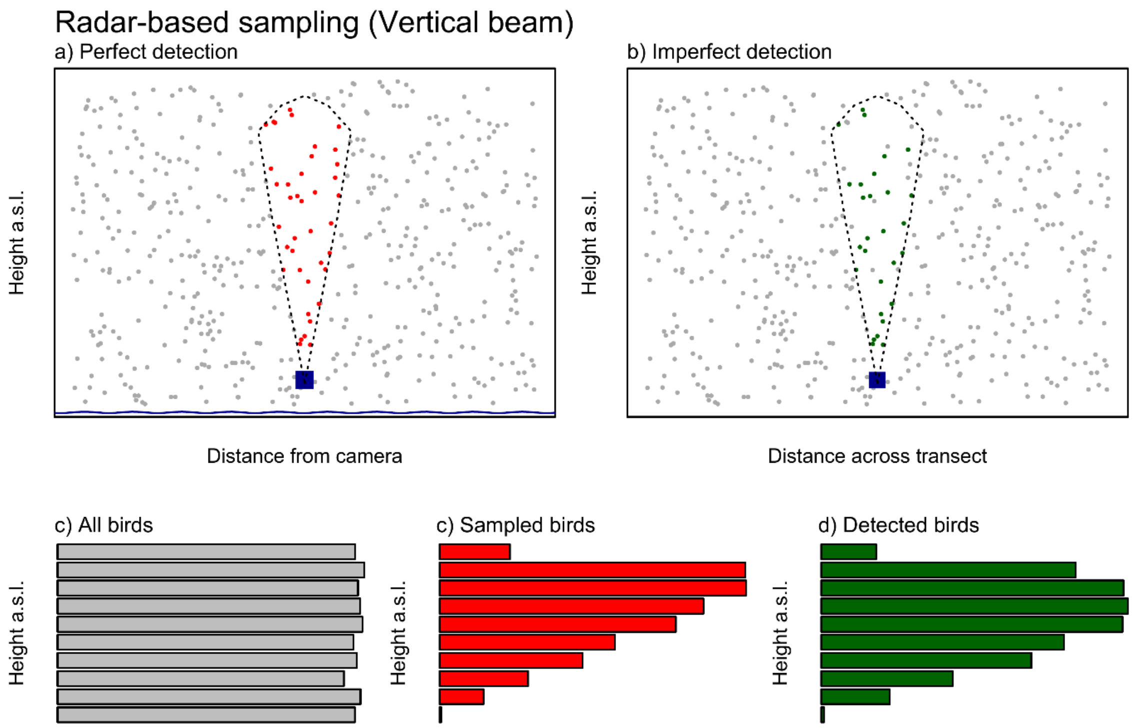
Marine radars traditionally rotate on a horizontal plane to detect collision hazards (e.g. ships, land) and provide information (e.g. bearing, distance) for avoidance. When set to rotate on a vertical plane they provide an altitudinal distribution of objects (e.g. birds) that pass through the beam (Figure 3.4). The elevation angle at which marine radar scan is not fixed but vertical (90°) orientations are typically used to optimise the vertical range (< 2 km). Marine radars are readily available (i.e. off the shelf) and as such are frequently used to sample bird flight height distributions in relation to human-made (e.g. transmission lines, wind turbines, communication towers) structures (Brabant *et al.* 2021, Bruderer *et al.* 2018, Fijn *et al.* 2015, Hilgerloh 2023, Hüppop *et al.* 2006).

Figure 3.4 Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for rotating vertical radar-based sampling methods. Scatterplots depict the sampled volume (cross section) which increases with distance from the sensor and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



Dedicated bird radars (e.g. BirdScan, Merlin, Robin, Accipiter, Birdtrack, ORJIP 2022) are specialised systems (hardware and software) designed specifically to monitor local (e.g. collision risk, migration) bird movements (Aschwanden *et al.* 2020, 2018, Pavón-Jordán *et al.* 2020). There are various operational setups (e.g. paired horizontal and vertical rotating radar, vertical pulse radar, phased array radar), each with different scanning elevation angles (e.g. 0–180°), operating ranges (vertical < 3 km, horizontal < 15 km), filtering options (e.g. rainfall, high waves) and detection capabilities (e.g. wing flap pattern, three-dimensional tracking). Wing flap patterns are used to distinguish between bird and non-bird (e.g. insect, bat) and further categorise birds (e.g. passerines, waders, large birds) and group size (e.g. individual, flock, Schmid *et al.* 2019, Zaugg *et al.* 2008).

Figure 3.5 Schematic illustration of theoretical sampling volume and associated variation in sampling efficiency with height above sea level (a.s.l.) for pulsing vertical radar-based sampling methods. Scatterplots depict the sampled volume (cross section) which increases with distance from the sensor and is plotted with a) perfect and b) imperfect (decreasing with increasing distance from observers) detection of uniformly distributed points (i.e. birds). Histograms show the resulting sampling efficiency under d) perfect detection and e) imperfect detection. Grey points represent unsampled birds, red points represent all birds within the sampled volume and green points represent all detected birds within the sampled volume. Grey bars represent a simulated uniform distribution of birds across height, red bars show the apparent flight height distribution arising from not adjusting for the sampled volume and green bars show the distribution from not adjusting for sampled volume and a variable detection probability.



Sampling characteristics

Radar systems record flight heights within predetermined locations and therefore use a Eulerian experimental design (Table 4.1). Radar systems can continually (i.e. 24 h d^{-1}) monitor the volume of air above (< 5 km) and surrounding (< 250 km) their location, and may generate hundreds of thousands to millions of observations depending on the setup (e.g. radar number, frequency, beam shape, scanning pattern/angle) and study duration (Fijn *et al.* 2015, Nilsson *et al.* 2018, Van Erp *et al.* 2023). They typically operate from near to ground or sea level and can be stationary (e.g. turbines, building, tower) or mobile (e.g. vessel, trailer). Radars cannot identify individual birds to species level and therefore do not provide species-specific values of flight height. They are, however, routinely paired with additional data collection methods (e.g. camera-based systems, observers) which enable species identification and allows for additional (e.g. behaviour) data to be collected simultaneously (ORJIP 2022, Skov *et al.* 2018). Radars are non-invasive but may affect the behaviour of birds depending on the location (e.g. structure, vessel, aircraft) of the equipment.

The sampled volumes covered by radar systems generally have complex shapes depending on the instrumentation used, and any derived flight height or density distributions need to be corrected for this. The theoretical beam shape can generally be derived when device characteristics are known, and theoretical beam patterns in combination with radar measurements have been used to derive altitude-band specific correction factors (so called migration traffic rate factors or MTR-factors; Liechti *et al.* 2019, Schmid *et al.* 2019) to account for variation in the sampled volume with altitude. Empirical volume calibration has been conducted for a horizontal marine radar used for bird movement analysis (Urmy & Warren 2017). Despite beam-shape adjustments, cross-calibration studies have found substantial differences in estimated bird flux rates between different radar devices (e.g. (Liechti *et al.* 2019, Weisshaupt *et al.* 2023, 2017), suggesting that – in addition to spatial variation in true flux rates between radar locations – the beam-shape adjustments may need improvement, as well as that detection characteristics within the beam require further consideration, such as differences in the sensitivity and range of object detection or algorithmic differences in signal processing and data analysis (Urmy & Warren 2020, Weisshaupt *et al.* 2023).

The detection probability of radar systems is negatively related to distance from the equipment but positively related to object size (Dokter *et al.* 2013, May *et al.* 2017, Schmid *et al.* 2019). The surveyed volume therefore varies with object size (i.e. species) and the number of birds detected at greater distances and altitudes will potentially be relatively low (i.e. negative observation bias). Species-specific detection probabilities must be estimated, and data collection restricted to distances for which probability is reliable (e.g. > 80%). The detection probability additionally varies in relation to the aspect (e.g. broadside, head on, tail on) and speed of the target bird (McCann & Bell 2017, Urmy & Warren, 2017).

The effective detection range of radar has been assessed using UAVs, theoretical objects and by comparing detection rates with those from observers (Dokter *et al.* 2013, May *et al.* 2017, Phillips *et al.* 2018, van Erp *et al.* 2023).

The detection range of a dedicated bird radar (Merlin) was estimated using a UAV (wing span 2.1 m) fitted with a GPS device and assuming a threshold detection probability of 0.5 (May *et al.* 2017). The detection range of gulls, ducks and geese was estimated to be < 1.5 km while for swans it was < 2.0 km.

The detection range of another dedicated bird radar (RobinRadar) was assessed by comparing detection rates with those from visual surveys. The detection range (50% probability) was estimated to be < 1.5 km for a range of coastal birds (Dokter *et al.* 2013).

The detection range of another dedicated bird radar (RobinRadar 3D Fixed) was estimated using theoretical objects within radar cross sections representative of a Carrion Crow (*Corvus corone*) and a Song Thrush (*Turdus philomelos*, van Erp *et al.* 2023). The detection range (80% probability) was estimated to be altitudes of < 300 m for small birds (< 62.5 g) and < 600 m for larger (500 g) birds.

The detection capabilities of another dedicated bird radar (Accipiter) were examined by comparing detection rates with field observations of individual and flocks of birds (Phillips *et al.* 2018). Of all observed bird movements (n = 972), 15% were detected by the radar of which 12% were individuals and 17% were flocks. Most of the birds observed and tracked were medium to large species (e.g. Red-tailed Hawk, Canada Goose) and the detection range was < 4.8 km. Detection probability was observed to decrease with distance from the radar unit and performance was deemed best < 2 km.

Radar data include many well know sampling errors and observation biases, and require considerable processing before subsequent analyses (Tjørnløv *et al.* 2023, van Erp *et al.* 2024). The bird detection probability of radars is not uniform over the whole radar observation window and there is a minimum and maximum distance from the equipment at which birds can be reliably detected. For vertical radar equipment placement dictates the minimum observable altitude, potentially leaving a considerable ground-level/sea surface-level blind zone. For radars rotating in the vertical plane this can in principle be overcome, but interference may still lead to a non-negligible blind zone near the surface. At small distances the radar beam is powerful enough to reflect on many unwanted features (i.e. false positive) and with increasing distance the radars detection probability decreases (negative observation bias). Nearby features (e.g. turbine rotors, sea surface) can disrupt bird detection (i.e. false positive or negative) in the area around them or block the radar beam (i.e. false negative) and non-bird objects (e.g. ships, waves) can temporarily increase detection rates (i.e. false positive). Environmental conditions (e.g. wind, rainfall, waves) can temporarily increase clutter (i.e. false positive) or activate filtering software which reduces detection sensitivity (negative observation bias).

Measurement characteristics
No studies could be found that have assessed the accuracy and precision of flight heights collected via radar.

Analytical considerations
The survey volume for a bird of given size (expressed as RCS) is defined by the maximum detection distance and the RCS specific beam angle (Kreutzfeldt *et al.* 2020, Schmid *et al.* 2019). Sampled volumes need to be

determined for each employed device, and ideally calibrated, and the data processing algorithms need to be considered in any assessment of detection characteristics (Urmy & Warren 2020). Radar estimates flight heights relative to MSL and provides continuous flight height data, values may therefore need converting before use in CRMs and derived distributions can be input directly into both the basic and extended Band models. The underlying distribution and associated confidence intervals can also be statistically modelled but no examples of this could be found in this review.

Operational considerations
Radar is particularly suited for continuous monitoring at fixed locations, although deployment on mobile platforms is in principle possible for relatively compact devices. Radar does not collect species-specific data but can be paired with visual methods (e.g. camera, observers). Systems can operate 24 h d^{-1} (i.e. not restricted by weather or to daylight hours) and the structures on which radar systems are mounted can usually support additional data collection methods (e.g. observers, cameras) and equipment for gathering supplementary environmental data. Radars are generally restricted to intra- and post-construction phases of wind farm development. The challenge of collecting radar observations close to the surface may limit their application to determine the proportion of birds at collision risk.

3.7. Light Detection And Ranging (LiDAR)

Light Detection And Ranging (LiDAR) is an active remote sensing technique that records the three-dimensional location of objects using pulses of light. The return records are aggregated to create detailed three-dimensional maps (i.e. point cloud) of surface structure which can be geo-referenced when combined with high resolution GPS data. There are two types of LiDAR system — discrete point return and continuous waveform systems (Anderson *et al.* 2016, Lefsky *et al.* 2002, Vierling *et al.* 2008). Discrete return systems (currently used for flight height sampling) measure the time taken for a laser pulse to travel to a single object and are used to determine height. Continuous waveform systems record the range to multiple targets and provide more detailed spatial information but carry a higher data processing cost. LiDAR has been used to model the 3D ecological structure of terrestrial environments for many decades, at markedly differing spatial scales (e.g. organs, individuals, populations, ecosystems) and for a wide variety of ecosystem (e.g. woodland, wetland, agricultural, grassland, urban) types (Davies & Asner 2014, Guo *et al.* 2021, Wang & Menenti,2021). LiDAR is a relatively new approach for estimating seabird flight height distributions (Cook *et al.* 2018, NIRAS 2018, Wicikowski *et al.* 2022).

Sampling characteristics
LiDAR follows a Eulerian sampling design whereby the sensors are fitted to an aircraft which follows predetermined transects (constant ground speed) within the area of interest (Table 4.1). LiDAR surveys are non-invasive and aircraft are piloted at altitudes (e.g. 350—400 m) thought to minimise disturbance but, as with digital aerial surveys, threshold values are based on facilitating population counts (i.e. quantified disturbance as birds being prompted into flight) and individuals in flight will be closer to the aircraft (Cook *et al.* 2018, Thaxter & Burton 2009). Sample sizes are moderate (100s to 1,000s) but LiDAR systems do not identify birds to species level and are therefore paired with high resolution cameras which simultaneously collect aerial imagery. Flights are restricted to calm and dry weather conditions, and daylight hours. The volume of air sampled by LiDAR sensors is pyramidal with the apex (i.e. narrow part of the beam) located at the sensor (i.e. the aircraft). Birds at higher altitudes are therefore less efficiently sampled (i.e. negative bias) and the resulting flight height distributions are potentially biased towards low altitudes.

The detection probability of LiDAR theoretically decreases with body size due to the increased surface from which points can be reflected. Larger species (e.g. Northern Gannet *Morus bassanus*) have been observed to reflect a greater number of LiDAR points compared to relatively smaller ones (e.g. Kittiwake *Rissa tridactyla*) but detection rates for the smallest species (e.g. European Storm Petrel) are currently not known (Cook *et al.* 2018). The former study did detect considerably greater numbers of larger species (i.e. gull sp.) relative to small species (e.g. auk sp.) but the position (e.g. broadside, tail-on) and behaviour (e.g. plunge diving, shearing) of individuals is also likely to influence detection probability. Detection probability is also expected to decrease with increasing distance from the sensor, either as the spread of laser pulses expands (reducing point density) or the power of return signals becomes too low to detect. The number of LiDAR points reflected from birds however has been reported not to vary (obviously) in relation to vertical (i.e. flight height) distance from the sensor or horizontal distance from the transect line (Cook *et al*. 2018). The study reported a point density of approximately 11 m^{-2} on the sea surface which is therefore likely suitable for detecting larger species. Birds towards the edge of the sampled volume are also less likely to be detected (i.e. false negative) and a reduction in the proportion of birds detected at distances greater than 125 m from the transect line supports this (Cook *et al*. 2018).

LiDAR data require processing before any subsequent analyses (Cook *et al*. 2018, Wicikowski *et al*. 2022). All points that potentially represent birds in flight must be detected and compared with aerial imagery for validation and, where possible, identification to species level. This can be achieved by first analysing (e.g. visual inspection, cluster analysis) the internal geometry (i.e. angle, distance) of the dataset to identify all potential birds. False positives generated at the sea surface (e.g. swell, water droplets) are commonly removed by setting a minimum flight height (e.g. 1–2 m) which can vary depending on the weather conditions but introduces negative bias towards low flight heights (Cook *et al*. 2018). The location of the remaining points (i.e. x, y, z) is then calculated using information from the LiDAR system and Inertial Measurement Unit (IMU) on board the aircraft before being compared with the aerial imagery (Cook *et al*. 2018). Alternatively, the location of all birds detected in the aerial imagery can be recorded and used to identify LiDAR data that correspond to their approximate location (Wicikowski *et al*. 2022). Both methods report that approximately 90% of birds could be matched in both LiDAR data and aerial imagery. There may, however, be a mismatch in sampled volumes, as well as a temporal lag between the LiDAR data and aerial imagery, both of which introduce uncertainty when attempting to match images and LiDAR points, particularly for large groups of birds at the sea surface.

Flight heights can be estimated by relating information from the LiDAR system to the altitude of the aircraft according to IMU (Wicikowski *et al*. 2022). This approach introduces error (vertical and horizontal) associated with estimating the aircrafts altitude into subsequent flight heights. An alternative approach is to compare all points classified as birds to those assumed to represent the sea surface (Cook *et al*. 2018). Flight heights (i.e. the difference) are then measured in relation to a reference geoid or the instantaneous sea surface height which generates a precise measurement of height above sea level that is independent from the aircraft's altitude.

Measurement characteristics

The accuracy of airborne LiDAR is typically assessed by comparing values with those from controlled sites. A series of ground check points are setup and the height of each is measured using a high precision GPS device (accuracy < 2–3 cm) and estimated using LiDAR (Cook *et al*. 2018, Wicikowski *et al*. 2022). The accuracy (mean error cm ± standard deviation) of a LiDAR survey aiming to estimate seabird flights in the Outer Forth and Tay Estuaries was assessed against the height of 12 ground control points (Cook *et al*. 2018). The average difference between the LiDAR measurements and those measured using GPS was 6 cm ± 2.5. The same study also attempted to measure accuracy by simultaneously flying three UAVs at heights of approximately 10, 40 and 80 m above sea level. The UAVs were overflown (n = 7) by an aircraft equipped with LiDAR. All three UAVs were clearly detected (2–9 points) in all flights over the study area and their flight heights estimated to within 100 cm (error range: -94–64 cm). The accuracy (RMSE) of a LiDAR survey aiming to collect seabird flight height data in the Moray Firth was measured using a grid of points spaced 50 cm apart over an area of land 500 cm x 500 cm. Three assessments were made which resulted in accuracies of 1.8 cm, 7.6 cm and 9.3 cm (Wicikowski *et al*. 2022).

Analytical considerations

LiDAR estimates flight altitude relative to MSL or flight height relative to the instantaneous SSH and values may therefore require conversion before use in CRMs. The method provides continuous flight height data and derived distributions and can therefore be input directly into both the basic and extended Band models. The underlying distribution can alternatively be statistically modelled, and confidence intervals estimated (Cook *et al*. 2018). A minimum sample size of 100–200 observations has been suggested necessary to fit robust distribution models but the effort required to achieve this may vary considerably (1–227 days) depending on the species and season of interest (Cook *et al*. 2018, Donovan & Caneco 2020). It is also not understood whether such relatively small sample sizes are representative of the variation in flight heights observed at the population level.

Operational considerations
LiDAR is particularly suited for near synoptic surveys of large areas. Aerial surveys are restricted to fine weather and daylight hours; the data are species-specific and can be collected for multiple species simultaneously. Additional behavioural data (e.g. flight speed) are embedded in the images, but no environmental data or individual characteristics can be simultaneously collected. The incorporation of two sensors (LiDAR and camera) requires an understanding of how the sampled volume of both devices varies with altitude. The aircraft must operate where the volume of both sensors is approximately the same for optimal results and there is likely a trade-off between survey coverage and detection probability which requires consideration (Cook *et al.* 2018). As with digital aerial surveys, aircraft must operate at higher altitudes when flown over structures (e.g. wind farms) and the change to sampling characteristics (e.g. survey volume, detection probability) will complicate any pre-post construction comparisons.

3.8. Animal-borne tracking devices
Telemetry-based measurements of flight heights can be collected through the deployment of animal-borne devices to track the three-dimensional movement of individual seabirds through space and time. Flight height is either estimated directly using trilateration (GPS and satellites) or indirectly via barometric altimeters (pressure sensors) which require calibration (Johnston *et al.* 2023, Schwemmer *et al.* 2023, van Erp *et al.* 2023).

Sampling characteristics
Animal-borne tracking data are Lagrangian by experimental design and are therefore (theoretically) unrestricted by study area boundary and can (theoretically) operate in all environmental conditions (i.e. weather, time of day, Table 4.1). In practice, spatial limitations are in part governed by the location of deployments, mediated by the movement ranges of the equipped animals, and there may be limitations in some environmental conditions (e.g. sampling rates and thus the precision of flight height estimates may be lower at night or in winter when batteries can't be so frequently charged). The number of individuals (i.e. sample size) that devices can be attached to is often limited (e.g. available animals, device cost) and not randomly selected (e.g. location, breeding status, age class); subsequent measurements are therefore potentially not representative of focal populations (Gibb *et al.* 2017, Watanuki *et al.* 2016). Animal-borne tracking devices are invasive and can negatively affect the physiology, reproduction, and survival of subject individuals depending on the attachment method and the relative device mass (Evans *et al.* 2020, Geen *et al.* 2019, Langlois Lopez *et al.* 2023, Seward *et al.* 2021). They can also induce a behavioural change (e.g. increased foraging, reduced commuting) which potentially limits their ability to accurately estimate natural flight height distributions (Gillies *et al.* 2020, Langlois Lopez *et al.* 2023, Longarini *et al.* 2023, Robinson & Jones 2014). In the UK, the use of tracking devices is licenced through the independent Special Methods Technical Panel of the national Ringing Scheme, to ensure that the welfare of study individuals is paramount, with requirements to assess potential impacts (typically through comparison with control samples). The relative device mass must be no greater than 3% to minimise impacts which limits their use to certain species and gaps in bird-borne data exist for many species and groups, and even below this body-mass threshold negative effects can occur (Langlois Lopez *et al.* 2023). Additional data can be collected while attaching the device (e.g. species, sex, biometrics) and throughout the individuals' subsequent activity (e.g. behavioural, environmental, physiological).

The devices deployed can generate tens to hundreds of thousands of flight height data points, but data collection is fundamentally limited by logistical considerations, including the availability of suitably trained and licensed fieldworkers and device constraints such as battery life. Trade-offs between high (i.e. near continuous sampling, typically < 20 s interval between fixes) and low (i.e. intermittent, typically > 1 minute interval between fixes) frequency sampling and consequently short or long deployment windows requires careful consideration (Clements *et al.* 2021). Researchers often attempt to optimise battery life by limiting high frequency sampling to periods when tags have surplus battery charge (i.e. during periods of sunlight) or are within specific areas (i.e. geofences, e.g. Johnston *et al.* 2023, Peschko *et al.* 2021, van Erp *et al.* 2023). While optimising battery life has many benefits (e.g. prolonged sampling period) such an approach may introduce positive bias towards sunny days (i.e. solar charging) and requires knowledge of where birds will travel. Measurement characteristics are additionally often not independent of sampling rates, low frequency sampling, for example, may not detect fine-scale behaviours (e.g. climbing to forage and plunge dive) with consequences for flight height distributions. Logistic constraints may further limit the ability to randomise samples across individuals and/or study sites, affecting the representativity of collected data with respect to the biological population.

Operational considerations
Spatial deployment of animal-borne methods is generally limited by access to suitable locations for safely and efficiently capturing and tagging individual animals, and are therefore often biased to seabird breeding colonies, which may or may not result in adequate observational coverage for areas of interest from a developer's perspective. There can be additional bias between or within colonies depending on ease of access. Telemetry-based methods are theoretically not restricted by environmental conditions (e.g. weather, time of day); the data are species-specific and may be collected for multiple species simultaneously (depending on device characteristics). Additional behavioural information, such as flight speed and turning angle (Akeresola *et al.* 2024), is embedded in the data, and environmental (e.g. temperature, air pressure, immersion, time/depth) and physiological (e.g. heart rate) sensors can be incorporated. Individual physical (e.g. mass, wing length), demographic (e.g. sex, age), biological (e.g. feather, faeces, blood), and behavioural (e.g. breeding, non-breeding) characteristics can be simultaneously collected while attaching the device.

3.8.1. GPS positioning

Flight height can be estimated via animal-borne tracking devices using trilateration to determine location based on data received from satellites (i.e. GPS). Three-dimensional positions are first calculated in relation to a mathematical model of the earth (e.g. geoid, ellipsoid) before digital elevation models (DEMs) are applied to estimate height relative to ground or sea level.

Measurement characteristics
There are several sources of uncertainty which are likely to influence the vertical accuracy of GPS trilateration (Péron *et al.* 2020, Poessel *et al.* 2018). Various models (i.e. geoid, ellipsoid) can be used to calculate an individual's three-dimensional position and tracking devices vary in which one is used (Péron *et al.* 2020). It is relatively simple to convert from one system of reference to another, but care must be taken as this represents a potential source of error in flight height data, and the specific reference system used by a device is not always easy to determine from device documentation (Péron *et al.* 2020). GPS locations contain horizontal errors that introduce uncertainty in the link to spatially explicit DEMs, particularly if topology is markedly variable (e.g. cliff edge). Errors in the original measurements from which DEMs are interpolated are assumed small (i.e. cm) relative to the other sources (e.g. LiDAR, Chai *et al.* 2022, Kim *et al.* 2022). The number of satellites used by the device to estimate position is also positively related to accuracy and more can be utilised as sampling frequency increases. Data are often filtered to limit analyses to locations produced above a threshold number of satellites (Lato *et al.* 2022, Schaub *et al.* 2023), although there is a trade-off between filtering data for quality control and explicitly modelling measurement errors. Filtering raw observations effectively truncates the observed flight height distribution, whereas explicit measurement error models may be able reconstruct the underlying true flight height distribution (Davies *et al.* 2024, Péron *et al.* 2020, Ross-Smith *et al.* 2016). Flight height data from animal-borne tracking devices are also likely to be spatially and temporally autocorrelated depending on the sampling interval (more likely with high frequency sampling). Autocorrelation within tracking data is commonly accounted for by filtering/resampling (e.g. larger time interval), averaging (e.g. individual, location, time), statistical modelling (e.g. autocorrelation structure, state-space-model) but it is also frequently ignored (Bogdanova *et al.* 2021, Borkenhagen *et al.* 2018, Johnston *et al.* 2023, Lane *et al.* 2020, Ross-Smith *et al.* 2016, Tarroux *et al.* 2016).

The vertical accuracy and precision of GPS trilateration has been estimated by comparing values with those of fixed stationary altitudes, drones and by identifying periods when birds were at known or measurable altitude (Acácio *et al.* 2022, Lato *et al.* 2022, Péron *et al.* 2020, Schaub *et al.* 2023, Thaxter *et al.* 2018, van Erp *et al.* 2023).

Comparisons of values (high frequency sampling) from two devices (CatLog Generation 2, Catnip Technologies, Hong Kong and OrniTrack-25, Ornitela, Lithuania) deployed on a UAV were made with measurements from a laser altimeter across a range of altitudes for stationary (2–50 m), horizontal (10 m) and vertical (2–60 m) flight patterns (Lato *et al.* 2022). The CatLog device significantly underestimated altitude for every stationary flight height (mean error = -54.93 m) and did so to greater an extent at lower (height range: 2–30 m; -33.96 m ± 27.15) flight heights compared to higher (height range: 40–50 m; -20.97 m ± 15.19). The OrniTrack-25 device also underestimated stationary flight height (mean error = -12.96 m), but differences were only significant at higher flight heights (height range: 20–50 m; -20.43 m ± 15.06). Both devices significantly overestimated (CatLog, 40.96 ± 29.99 m; OrniTrack-25, 6.5 m ± 3.13) altitude during horizontal movement but while the OrniTrack-25 device significantly underestimated altitude (-3.20 m) during vertical movement the CatLog device significantly overestimated altitude (40.40 m).

Schaub *et al.* (2023) compared height estimates derived from GPS devices (eight device models), made at high and low sampling frequencies, deployed on four raptor species (Montagu's Harrier *Circus pygargus*, Hen Harrier *C. cyaneus*, Marsh Harrier *C. aeruginosus*, Red Kite *Milvus milvus*) when they were stationary and at ground level (i.e. height above ground was approximately zero). Vertical accuracy (error from true height) did not differ (median range: -3.8–4.3 m) between sampling frequencies (i.e. high, low) but precision (absolute error from measured height) was considerably more variable for low (absolute median range: 2.6–17.4 m, mean = 6.3 ± 4.6) frequency sampling compared to high (median absolute error range: 1.0–4.0 m, mean = 2.4 ± 1.0).

The vertical and horizontal accuracy (mean error) of GPS/GPRS tracking devices (1-, 20- and 60-minute sampling frequency) was also measured while devices were stationary and both before and while deployed on pre-fledging White Storks *Ciconia ciconia* (Acácio *et al.* 2022). Both accuracy (horizontal: 1 min = 3.40 m ± 3.10, 20 min = 4.23 ± 4.15, 60 min = 6.50 m ± 8.34, vertical: 1 min = 4.95 m ± 4.12, 30 min = 6.56 ± 6.72, 60 min = 9.69 ± 19.28) and precision (horizontal: 1 min = 4.93 m ± 4.15, 20 min = 6.14 ± 5.46, 60 min = 9.15 m ± 9.46, vertical: 1 min = 3.60 m ± 5.94, 30 min = 8.79 ± 9.17, 60 min = 14.31 ± 24.95) increased with sampling frequency during stationary tests. Horizontal accuracy did not change after deployment (20 minute sampling frequency) of the devices on white storks (before: mean = 4.21 m ± 18, after: mean = 4.10 m ± 15) and vertical accuracy improved (before: mean = 7 m ± 71, after: mean = 6 m ± 56 m).

The vertical and horizontal accuracy (mean error with 0.05 and 0.95 quantiles) of GPS tracking devices was assessed (6, 60 and 600 second sampling frequency) using stationary tests and while attached to birds (White Stork and Honey-buzzard *Pernis apivorus*) at known heights (nest sites) above ground (Bouten *et al.* 2013). Both horizontal (6 s: mean = 1.13 m, 0.20, 2.33, 60 s: mean = 3.23 m, 0.63, 7.48, 600 s: mean = 29.95 m, 9.26, 108.10) and vertical (6 s: mean = 1.42 m, 0.25, 3.75, 60 s: mean = 3.99 m, 0.23, 9.76, 600 s: mean = 26.27 m, 2.13, 102.00) accuracy increased with increasing sampling frequency during stationary tests. Accuracy decreased slightly for the device (6 s sampling frequency) attached to the White Stork (horizontal: mean = 2.45 m, 0.34, 7.14, vertical: mean = 2.77 m, 0.38, 7.61) compared to stationary tests and considerably for the device (600 s sampling frequency) attached to the Honey-buzzard (horizontal: mean = 67.43 m, 14.65, 226.30, vertical: mean = 20.79, 0.76, 45.24).

The accuracy and precision of GPS tracking devices relative to MSL is regularly assessed by identifying periods when animals (particularly seabirds) are assumed to be at the sea surface (i.e. MSL, Johnston *et al.* 2023, Thaxter *et al.* 2018, van Erp *et al.* 2023). Tracking devices (UvA-BiTS) fitted to Lesser Black-backed Gulls *Larus Fuscus* in the Netherlands recorded a mean altitude of -4.36 m when locations (n = 99,137) were assumed to be at the sea surface. (van Erp *et al.* 2023). Similar devices (UvA-BiTS) fitted to Lesser Black-backed Gulls at two sites in the UK (Havergate Island, Isle of May) recorded median altitudes of -3 m and 0 m for Havergate Island and the Isle of May respectively (Johnston *et al.* 2023). The accuracy of altitude measurements from GPS devices attached to Lesser Black-backed Gulls (n = 2) was assessed for different sampling frequencies (10, 16, 60 and 300 seconds, Thaxter *et al.* 2018). Assessments were restricted to within two hours of mid tide and resulted in small variation in median altitude between sampling frequencies (median altitude of 3 m and 2 m to 2 m and 2 m between 10- and 300-second sampling frequencies respectively). Boxplot statistics showed the whisker limit range increased with increasing sampling frequency (11 and 14 m to 24 and 19 m between 10- and 300-second sampling frequencies respectively).

Analytical considerations

GPS positions estimate flight altitude relative to MSL and values may need converting before use in CRMs. The approach provides continuous flight height data and distributions can either be input directly into both the basic and extended Band models. The use of measurement error models to account for the typically noisy flight height measurement of low-frequency sampling or autocorrelation within high-frequency is recommended (Péron *et al.* 2020, Ross-Smith *et al.* 2016). The number of individuals required to determine two-dimensional (i.e. area use) interactions between seabirds and offshore developments has been estimated for a number of species (Guillemot *Uria aalge* > 19, Razorbill *Alca torda* > 3, Atlantic Puffin *Fratercula arctica* > 20, Kittiwake *Rissa tridactyla* > 11–20 (site-specific), Lesser Black-backed Gull *Larus fuscus* > 13–41, Bogdanova *et al.* 2021, Thaxter *et al.* 2017). The sample size required to characterise three-dimensional space use (i.e. flight height distribution) has not been formally assessed but is potentially both species and site specific (Bogdanova *et al.* 2021, Lascelles *et al.* 2016, Soanes *et al.* 2015, Thaxter *et al.* 2017).

3.8.2. Barometric altimetry

Barometric pressure sensors can be attached to tracking devices and flight height estimated by measuring atmospheric pressure at each position and comparing the values with those at sea level using the barometric formular (Johnston *et al.* 2023, Lane *et al.* 2020, Péron *et al.* 2020).

Measurement characteristics

The relationship between altitude and pressure is responsive to changes in the weather, and the barometric formula ideally requires an additional measure of sea/ground level pressure for every location (i.e. calibration). This is rarely feasible in practice, and measures of sea level pressure are generally estimated either using values from additional devices fixed at sea/ground level, from the periods when the attached device is predicted to be at sea/ground level (e.g. bird resting on the surface) or from modelled remotely sensed data (Cleasby *et al.* 2015, Johnston *et al.* 2023, Lane *et al.* 2020). The consequence is that barometric altimeters are commonly not calibrated using in situ observations of sea level pressure, but values are modelled at coarser spatial grain than the animal movements which generates temporal autocorrelation in the error time series and a systematic over- or underestimation of flight height (Lato *et al.* 2022, Péron *et al.* 2020, Schaub *et al.* 2023). The vertical accuracy and precision of barometric altimetry has been estimated by comparing values with those from a drone and by using periods when birds fitted with tracking devices were at known heights (Lato *et al.* 2022, Schaub *et al.* 2023). As with GPS positioning, data are likely to be temporally or spatially autocorrelated which can be addressed in several ways (discussed above).

Comparisons (5-second sampling frequency) of values from a device (AxyAir) with a UAV (fitted with a laser altimeter) were made across a range of altitudes for stationary (2–50 m), horizontal (10 m) and vertical (2–60 m) flight patterns (Lato *et al.* 2022). Accuracy (mean error ± standard deviation) did not vary across flight heights but improved from stationary (-2.14 m ± 11.63) to moving horizontal (-1.0 m ± 3.35) but not vertical (-2.0 m) movement. While no significant difference in measurements from barometers and laser altimeters was observed, the barometer consistently underestimated altitude.

Schaub *et al.* (2023) also compared height estimates derived from devices with barometers (three device models), made at continuous and > 5 minute sampling frequencies, deployed on four raptor species (Montagu's Harrier, Hen Harrier, Marsh Harrier, Red Kite) when they were stationary and at ground level (i.e. height above ground was approximately zero). Both vertical accuracy (error from true height) and precision (error from measure height) did not differ between low and high sampling frequencies. Height above ground was consistently underestimated (median range: -15 – -4.9 m) but precision for both low (median absolute error range: 2.8–4.2, mean = 3.5 ± 0.7) and high (median absolute error range: 2.3–3.5 m, mean = 2.9 ± 0.6) frequency sampling was comparable to that of high frequency sampling via trilateration from the same study.

Analytical considerations

Barometric altimetry estimates flight height relative to the sea surface and values need converting to MSL before use in CRMs. The approach provides continuous flight height data and distributions can either be input directly into both the basic and extended Band models or statistically modelled (e.g. space-state models) to account for measurement error and erroneous records, autocorrelation, and provide confidence intervals (Péron *et al.* 2020, Ross-Smith *et al.* 2016). Appropriate sample sizes are the same as for GPS positioning (discussed above).

4. DISCUSSION

4.1. Strengths, weaknesses, opportunities, and limitations

Many methods have been developed for estimating seabird flight height distributions and a number of significant reviews have assessed and compared them (e.g. Largey *et al.* 2021, Nilsson *et al.* 2018, Péron *et al.* 2020). Such studies have improved current understanding of each method's limitations and how the uncertainty associated with the flight height estimates they produce and thus collision risk estimates might be reduced. Table 4.1 provides a summary of the characteristics of the methods and associated analytical operational considerations that have been identified within this review.

The measurement accuracy and precision of flight height estimates produced by many existing methods (e.g. rangefinders, LiDAR, high frequency GPS) is relatively high (< 10 m), and advanced statistical techniques (e.g. state-space model, nonlinear models) have been developed that can account for error by describing the underlying distributions and providing confidence estimates (Cook *et al.* 2018, Johnston *et al.* 2014, Johnston & Cook 2016, Ross-Smith *et al.* 2016). Such measurement error models are however generally only applicable when the expected errors are small compared to the desired scale of inference (approximately 10 m when collision risks are considered). Several of the discussed methods (i.e. low frequency GPS devices, single-camera photogrammetry) have expected measurement errors so large (> 50 m) that measurement error models may not be able to meaningfully recover underlying flight height distributions. Such methods are not currently recommended for estimating seabird flight height distribution.

Sampling uncertainty, however, is generally much less well understood and is often not incorporated into analytical workflows. This may be an important reason, why — despite significant efforts — there remains a lack of agreement in the flight height estimates produced by different methods, with a resultant degrading of confidence in collision risk estimates (Borkenhagen *et al.* 2018, Johnston *et al.* 2014, Johnston & Cook 2016). Not only may different methods result in different flight height distributions for given species, but there can also be differences between studies using the same methods at different sites, reflecting spatial, temporal, and individual variation. The ad hoc way observational studies and calibration/validation experiments are often designed is a key driver of uncertainty. A lack of transparency (e.g. language, methods, validation) in published results is slowing progress and environmental advisors are increasingly reliant on data providers to explain the use of novel technology. There is also limited access to raw sensor data (due to commercial and intellectual property rights) which in turn limits subsequent analyses and assessment of observation technologies.

Nevertheless, technological advances are generating a wide range of new opportunities for flight height studies, and such advances are necessary if flight height monitoring is to be scaled out in spatial and/or temporal coverage for development and/or operational purposes. Manufacturers of tracking devices, for example, are providing increasingly complex programming options which are required for designing robust experiments and additional sensors which facilitate more intuitive switching between high and low frequency sampling (e.g. high-frequency tracking only when the bird is in flight, Harel *et al.* 2016). Continuous waveform LiDAR systems can distinguish between surface type (e.g. water, land, vegetation) which might remove uncertainty associated with birds < 2 m from the sea surface (Wang & Menenti 2021). Moving forward, it is essential that all equipment (e.g. sensor, GPS) specifications (e.g. precision, accuracy), calibration (e.g. sample volume, detection probability) and processing (e.g. filtering, supplementary data) methods, and raw data (e.g. sensor, aircraft GPS) are routinely made publicly available for new sampling technologies. Best practice guidance must be developed (and followed) and frequently updated to keep pace with current understanding.

4.2. Operational considerations

Logistical (e.g. installation of equipment on offshore turbines) and economical limitations potentially restrict each method to certain scales (spatial and temporal) of monitoring.

Table 4.1 **Summary of the characteristics (sampling and measurement) and considerations (analytical and operational) associated with methods for sampling the flight height distributions of seabirds. The table summarises each methods approximate sample size and whether each can identify individuals to species level, sample during nocturnal hours or poor weather and whether each is invasive (e.g. potentially harmful) or may disturb individuals (e.g. invoke behavioural responses). The approximate vertical range (m) and whether sampling is potentially biased (positive + or negative -) towards low or high altitudes is also provided. Approximate accuracy is given where known, as is whether the detection probability (and therefore sampled volume) can be currently determined. The value (FH – Flight height relative to the instantaneous sea surface, FA – Flight altitude relative to a reference sea level, e.g. MSL) obtained, the type of data generated (discrete ordinal bands, continuous), whether error models can be utilised and if any post processing of data is required is given. Finally, the scale each method currently operates at and the construction phase at which each can be deployed is provided. Information is given by sampling method and design (Eulerian, Lagrangian).**

	Sampling characteristics								Measurement characteristics		Analytical considerations				Operational considerations	
Method	**Design**	**Sample size**	**Species-specific**	**Night**	**Rain / wind**	**Invasive / disturbance**	**Potential vertical range**	**Observation bias**	**Accuracy / precision**	**Detection probability**	**Value**	**Data type**	**Error model**	**Post processing**	**Scale**	**Phase**
Visual survey	Eulerian	103	✓	×	×	✓	< 300 m	+ low / - high	Unknown	Part known	FH	Ordinal	✓	✓	Regional	pre / intra / post
Rangefinder	Eulerian	103	✓	×	×	✓	< 100 m	+ low / - high	± 10 m	Unknown	FH	Continuous	✓	✓	Regional	pre / intra / post
Mono photogrammetry	Eulerian	103	✓	×	×	Unknown	< 500 m	- low / - high	± 100 m	Unknown	FA	Continuous	×	✓	Regional	pre / intra / post
Stereo photogrammetry	Eulerian	Unknown	✓	×	×	×	< 300 m	- low	± 1 m	Unknown	FA	Continuous	✓	✓	Local	intra / post
Microphone array	Eulerian	Unknown	✓	✓	×	×	< 150 m	+ low / - high	± 10 m	Part known	FA	Continuous	✓	✓	Local	intra / post
Radar	Eulerian	106	×	✓	×	×	< 2000 m	+ low / - high	Unknown	Known	FA	Continuous	✓	✓	Local	pre / intra / post
LiDAR / camera	Eulerian	103	✓	×	×	Unknown	< 500 m	+ low / - high	± 1 m	Unknown	FA or FH	Continuous	✓	✓	Regional	pre / intra / post
Bird-borne GPS	Lagrangian	10	✓	✓	✓	✓			± 10 m	NA	FA	Continuous	✓	✓	Global	pre / intra / post
Bird-borne barometer	Lagrangian	10	✓	✓	✓	✓			± 10 m	NA	FH	Continuous	✓	✓	Global	pre / intra / post

Observer-based methods (e.g. visual surveys, rangefinder) can be deployed at various spatial (e.g. local, regional, national) and temporal (e.g. days, weeks, months) scales but are limited to fine weather and daylight hours. They can occur pre-, intra-, and post-construction of wind farms and have no long-term maintenance requirements and minimum post-processing requirements (e.g. conversion to MSL). Their cost incorporates the requirements of observers (e.g. training, safety, living), their equipment (e.g. optics, rangefinders, PPE) and the survey vehicle (e.g. fuel, crew), which can be considerable.

Sea-level camera-based methods (e.g. stereophotogrammetry) are limited to local spatial scales but can operate over various temporal (e.g. days, weeks, months) scales. They are however typically limited to daylight hours, and detection ranges vary with weather conditions. Cameras require structures on which they can be mounted and are therefore restricted to intra- and post-construction, all equipment requires maintenance (hardware and software) while in operation and there are considerable post-processing requirements (i.e. image analysis). The cost of sea-level camera-based method includes the purchase, installation, calibration, and ongoing maintenance of equipment/post processing of data by trained engineers.

Microphone-based methods (e.g. arrays) are limited to local spatial scales but can operate over various temporal (e.g. days, weeks, months) scales. They are limited to fine weather but can operate in both daylight and nocturnal hours. They require structures on which they can be mounted and are therefore restricted to intra- and post-construction. Devices also require maintenance (hardware and software) while in operation and significant post-processing once data are collected. The cost of microphone-based methods includes the purchase, installation, calibration, and ongoing maintenance of equipment/post-processing of data by trained engineers.

Radar-based methods are typically limited to local spatial scales, but their range can be extended to regional or national scales via the construction of networks. They can operate at various temporal scales but require structures (i.e. restricted to intra- and post-construction unless platforms are present), maintenance (hardware and software) while in operation and considerable post processing of data. Radar can operate continuously (i.e. 24 h d^{-1}) but may experience interference when environmental conditions (e.g. wind, rain) are moderate to severe. The cost of radar-based methods includes the purchase, installation, calibration, and ongoing maintenance of equipment/post processing of data by trained engineers.

Aircraft-based methods (e.g. LiDAR, aerial imagery) typically operate at local or regional spatial scales and can potentially be extended to national scale by increasing the number of survey vehicles. They are restricted to small temporal scales (e.g. hours) but this may also be expanded by increasing the sampling frequency (e.g. daily, weekly, monthly). Aircraft-based methods are restricted to fine weather and daylight hours (both LiDAR and DAS can potentially be deployed at night, but species ID is not possible), they can be used pre-, intra-, and post-construction of wind farms and have no long-term maintenance requirements. Aircraft are however restricted to higher altitudes where piloted over wind farms due to safety regulations They can also operate far offshore where the use of other methods may not be feasible. The cost of aircraft-based method includes the purchase, installation and calibration of equipment, operation of the aircraft (e.g. pilot, fuel) and post-processing of data.

Telemetry-based methods operate across all spatial scales (e.g. local, regional, national, global) but are typically restricted to smaller temporal scales (e.g. weeks, months) due to device and attachment limitations (e.g. battery, tail-mounted devices). They can be used during pre-, intra-, and post-construction of wind farms and have no long-term maintenance requirements. The cost of telemetry-based method includes the purchase, programming, and attachment (e.g. licenses, PPE) of devices.

4.3. Best practice recommendations

There should be some basic minimum standards that flight height sampling methods must meet if they are to generate robust estimates of collision risk with wind turbines via CRMs. These might be that data are species-specific and representative of all annual cycle stages (i.e. breeding, non-breeding, migration) that can be encountered within the broad area of interest. Sample size, accuracy and precision should be sufficient (here ± 10 m) such that underlying flight height distributions can be statistically modelled (e.g. Cook *et al.* 2018, Ross-Smith *et al.* 2016). Methods must also ideally be capable of both sampling prior to wind farm construction (i.e. baseline data collection) and being operated at regional/national scales (Largey *et al.* 2021, Searle *et al.* 2023). Rangefinders, LiDAR, and animal-borne tracking devices (high-frequency GPS) broadly meet these requirements or can be incorporated into experimental designs that do so (Table 4.1).

The following principles are suggested for all flight height studies looking to generate data to inform flight height distributions for use in CRMs.

- Studies must make considerable efforts to sample across space (e.g. habitat), time (e.g. day, night, weather), and individuals (e.g. sex, age, behaviour) in an unbiased way (i.e. random sampling).
- Appropriate sample sizes should be determined via power analyses to characterise individual or population level flight height distributions at the desired temporal and/or spatial resolutions.
- Clear and repeatable descriptions of all the methods used to collect and process data should always be provided (e.g. MSL model, species specific detection probability, effectively sampled volume, work flow) to allow for comparison and assessment by others (e.g. van Erp *et al*. 2024).
- Many devices (e.g. rangefinders) provide best practice guidance (e.g. calibration, operation, storage) which should be reported and followed.
- Simultaneous environmental (e.g. wind, sea surface height, tidal phase) and behavioural (e.g. foraging, commuting) characteristics should also be routinely collected using standard approaches where possible.
- Data should be reported in a standard format (e.g. flight height relative to MSL ± 95% confidence intervals), using common language, and along with all relevant metadata (e.g. environmental conditions, SSH, behaviour, platform/observer height) submitted to open data repositories for future meta-analyses.
- Rigorous quality assurance and control procedures should be outlined and implemented throughout the collecting and processing of data.

Routine assessments of the accuracy and precision of methods are not currently common practice but are fundamental to understanding (i.e. measurement characteristics) and adequately communicating uncertainty (limitations, considerations) in observed flight height distributions (Searle *et al*. 2023). This is particularly important where new or untested equipment is employed. Such assessments should be undertaken in both experimental and field conditions, and prioritised over simply collecting additional data (e.g. Harwood *et al*. 2018, May *et al*. 2017, Schaub *et al*. 2023). Experimental designs that facilitate the assessment of three-dimensional accuracy and precision are required to fully understand the limitations of different methods. This can be achieved by comparing values with those from moving (e.g. UAVs, balloons, kites) or stationary (e.g. ground control sites, structures) targets of known size and position. The accuracy and precision of animal-borne tracking devices varies considerable depending on the model, measurement method (i.e. trilateration, barometric altimetry, sampling frequency), environmental conditions and animal behaviour (e.g. mobile, stationary). Routinely testing their accuracy and precision both pre- (e.g. stationary test, UAV) and post- (e.g. roosting, nesting, floating) deployment should be a basic requirement (Acácio *et al*. 2022, Lato *et al*. 2022, Schaub *et al*. 2023).

The flight height data collected using the methods discussed provide an imperfect sample of the true underlying distribution, and it is increasingly viewed as best practice to statistically model both the sampling and measurement characteristics that are thought to influence how the observed distribution arises from the true underlying distribution (Péron *et al*. 2020, Searle *et al*. 2023). A range of formal error models have been assessed for some of these aspects and all should be considered to avoid overconfidence in the results (Cook *et al*. 2018, Johnston *et al*. 2014, Johnston & Cook 2016, Ross-Smith *et al*. 2016). More complex analyses of flight height distributions will likely require increasingly interdisciplinary work (e.g. animal behaviour, statistical analysis, computer programming) but all subsequent inferences are expected to be more reliable.

Estimating the effectively sampled volume and accounting for heterogeneous detection (vertical and horizontal dimensions) is essential to Eulerian sampling methods and should be a routine component of flight height assessments. Methods for which the maximum sample volume can be theoretical determined (e.g. LiDAR, aerial imagery) should provide the value so flight height distributions can be adjusted accordingly. Aircraft-based surveys should ideally employ Global Navigation Satellite Systems (GNSS) and Inertial Measurement Units (IMU) which, combined with detailed knowledge of the camera and LiDAR setup, would allow accurate quantification of the sampled airspace during a survey. Novel analytical methods need developing before detection probabilities can be formally assessed but data that describe the three-dimensional locations of birds in relation to the sensor/observer will be required and should be routinely collected. This can be achieved via visual surveys by recording flight height, and both distance and bearing from the observer, or if observers are equipped with rangefinders, elevation angle/flight height, azimuth/bearing (true), and distance to target/horizontal distance. Digital aerial surveys can record flight height and spatial coordinates (via georeferenced images) or if equipped with LiDAR, both the duration and angle of the return signal.

4.4. Knowledge gaps and research priorities

Knowledge gap 1
There is currently a general lack of analytical methods to account for the heterogeneous detectability that is inherent to all Eulerian data collection methods (Table 4.1). A comprehensive understanding of the species-specific detection probabilities associated with each method is required to determine the effectively sampled surveyed volume. The effectively sampled volume is required to accurately estimate height frequency distributions but is currently unknown for most if not all Eulerian methods and particularly so for observer-based (e.g. rangefinder) approaches. The development of protocols and analytical methods for determining species-specific detection probabilities of Eulerian sampling methods is therefore a priority. This additionally offers the potential to further improve the accuracy of abundance estimates derived from equivalent surveys.

Quantifying variation in detection probability for observer-based sampling methods (e.g. rangefinders) will require the development of novel statistical approaches. As previously noted, data that describes the three-dimensional position of birds in relation to observers will be required and the accuracy and precision with which it is collected must be understood. Both sampling procedures and experiments must therefore be designed to further understanding in this field.

Variation in the detection probability of discrete LiDAR sampling can also be quantified by examining relationships between the characteristics of targets (e.g. size, distance, colour) and accompanying return signals (e.g. number, angle). The number of points reflected decreases with target size and a theoretical minimum detection size can potentially be determined (Cook *et al*. 2018). The number of points reflected by targets should theoretically also decrease with increasing distance from the sensor due to beam divergence. While this is not reported, fewer targets have been detected at the edge of LiDAR transects which supports this to some degree and further investigation is required (Cook *et al*. 2018).

The detection probability of discrete LiDAR systems is also influenced by atmospheric conditions (e.g. moisture) and sea surface attributes (e.g. waves, spray). The sea surface is particularly problematic because abundant false detections (combined with the lag between LiDAR and imagery systems) creates uncertainty when attempting to match imaged birds with LiDAR points. Efforts must be made to minimise the subsequent accepting of false positive or truncation of data. For example, the sea surface is currently assumed a smooth plane (i.e. MSL) which ignores the presence of swell. The surface is in fact irregular and treating it so (e.g. topographic model, Varbla *et al*. 2021) would be more reflective of the natural system. LiDAR can also be used to simultaneously map the instantaneous sea surface and flights height measured relative to the result. Subsequent estimates of flight height would consequently vary depending on each birds location (i.e. wave or trough), the application of such data to CRM however needs to be investigated further. Alternatively, continuous waveform LiDAR systems are not hindered by atmospheric interference and provide additional information regarding reflectivity (i.e. surface properties) of the target. Such systems may remove uncertainty associated with the sea as they allow for the classification (and subsequent removal) of sea surface characteristics.

Knowledge gap 2
The extent to which the behaviour of birds in flight is altered by different sampling methods is not well understood, particularly for telemetry- and aircraft-based methods. The flight characteristics (e.g. height, speed) of seabirds are strongly related to behaviour (e.g. foraging, commuting) and flight height distributions are therefore potentially sensitive to sampling methods that induce behavioural change (Fijn & Collier 2022, Ross-Smith *et al*. 2016, van Erp *et al*. 2023).

Animal-borne tracking devices can significantly alter the behaviour (e.g. increased foraging) of individual seabirds such that it may not be representative of the focal population (Bodey *et al.* 2018, Geen *et al.* 2019, Gillies *et al.* 2020, Longarini *et al.* 2023). Survey vessels may also induce a behavioural response (i.e. attraction, disturbance) from seabirds which may affect subsequent flight height distributions (Jarrett *et al.* 2021, Mendel *et al.* 2019, Schwemmer *et al.* 2011). While some information is available regarding the response of seabirds at the sea surface to aircraft (e.g. Thaxter & Burton 2009), no information could be found regarding the response of seabirds in flight. A better understanding of how sampling methods influence individual behaviour, and the consequences in terms of flight height distributions is needed.

Knowledge gap 3

Few methods sample across time in a non-biased manner and there is subsequently a limited quantity of flight height data that incorporates potential behavioural responses to the complete range of environmental conditions (e.g. strong winds, nocturnal) experienced by sea birds (Table 4.1). Sampling across environmental conditions as they naturally occur is fundamental to determining true flight height distributions and remains a key challenge for most sampling methods.

Animal-borne tracking devices can theoretically sample across time in a non-biased manner and are therefore well placed to address this challenge, although studies may be limited in the numbers of sites and individuals sampled and thus their spatial representation. An increased/renewed focus within telemetry studies on quantifying species-specific relationships (temporal and spatial) between environmental conditions (e.g. wind speed, wind direction, temperature, diel cycle phase) and flight characteristics (e.g. height, speed, distance) is therefore required (Davies *et al.* 2024, Kumagai *et al.* 2023, Ross-Smith *et al.* 2016, Tarroux *et al.* 2016, van Erp *et al.* 2023). The sampling characteristics (e.g. sample size, tracking duration) required to make robust three-dimensional population level inferences (i.e. flight height distribution) have however not been formally determined and additional assessments are needed (e.g. Lascelles *et al.* 2016, Soanes *et al.* 2015, Thaxter *et al.* 2017).

While most Eulerian sampling methods (i.e. rangefinder, LiDAR) are temporally biased towards particular conditions (e.g. calm weather, daylight) they are able to collect environmental (e.g. wind speed, wind direction, temperature) and behavioural (e.g. flight height, flight direction, flight type) data for multiple species simultaneously (Ainley *et al.* 2015, Linder *et al.* 2022). Eulerian sampling methods can also be designed to operate at annual temporal scales thereby encompassing all species-specific annual cycle stages (i.e. breeding, non-breeding, migration) that can be encountered (Ainley *et al.* 2015, Linder *et al.* 2022).

Understanding and quantifying the functional responses of seabirds to variation in environmental conditions will ultimately improve understanding of current flight height distributions and facilitate predictions of flight height in conditions not currently represented but of particular interest with regards to collision risk (e.g. fog, storms, night). Functional responses are also fundamental to additional modelling approaches (e.g. Agent-Based Models, ABMs) which are playing an increasingly important role in predicting the response of seabird populations to wind farm developments (Stillman *et al.* 2015, van Bemmelen *et al.* 2021, Warwick-Evans *et al.* 2018).

4.5. ReSCUE project research priorities

To address some key knowledge gaps discussed in the report, a large part of the proposed research under the ReSCUE project will focus on improving current understanding of the accuracy, precision, and detection probabilities of observations from airborne platforms, particularly LiDAR. Both experimental trials and gap-filling surveys of seabirds should be considered in which targets of various characteristics (e.g. size, orientation, colour, surface type), heights and distances (relative to the sensor/observer) are measured using LiDAR and imagery, ideally paired at the individual target level. Airborne measurements should be complemented, where logistically possible, by simultaneous observations from the ground, for example with rangefinders, and high-frequency tracking devices. However, independent investigations of accuracy and precision for these complementary methods could be conducted in isolation if necessary.

The assessment of accuracy, precision, and detection probabilities of continuously sampling sensor-based methods such as camera systems, radars, or low-frequency tracking devices, likely require substantially different experimental approaches, and we therefore consider the validation of such technologies to be outside of the scope of the ReSCUE project. This does not imply that such technologies cannot form an important contribution to a better understanding of flight height distributions.

5. CONCLUSIONS

The review has highlighted that no one method provides information that is representative of all environmental conditions or of spatial variation, for a given species. Thus, for the purposes of producing representative flight height distributions for use in CRMs, integration of information across multiple measurement methods is likely to be required to overcome the errors and biases inherent to all flight height data sampling methods.

Considering data from Lagrangian and Eulerian methods simultaneously often results in more balanced insights and reduced uncertainty when estimating the distribution of seabirds (Carroll *et al.* 2019, Fischer *et al.* 2023, Phillips *et al.* 2019, Sansom *et al.* 2018). The strengths and weaknesses of each sampling method can also become more apparent/informative when comparing data from different sampling approaches (Carroll *et al.* 2019, Phillips *et al.* 2019). The asymmetries and symmetries within comparisons improve understanding of how (e.g. where, when, effort) different sampling methods should be employed and help to prioritise data collection. It is therefore useful to undertake flight height sampling via multiple methods simultaneously within the same area. For comparisons to have value, both sampling and analyses must be designed specifically such that similarities and differences between the resulting outputs are informative.

Producing generic flight height distributions for CRMs will likely require combining data from multiple sampling methods, sources, and studies (Matthiopoulos *et al.* 2022). Integrating data from the different sampling approaches will have to go beyond simply pooling data from different approaches, or even different instrumentations of the same approach because of the inherent assumptions, constraints and biases associated with each method (Camphuysen *et al.* 2012, Watanuki *et al.* 2016). Common analytical challenges to data integration include data scale mismatches (i.e. dimension, resolution), unbalanced data (e.g. sample size, information content) and sampling biases (e.g. detection probability, detection accuracy). However, statistical frameworks and best practices for data integration in biodiversity monitoring have been developing in recent years, and a key component of these frameworks is an explicit accounting for observation characteristics of each contributing data source (Isaac *et al.* 2020, Mancini *et al.* 2022). This is not merely an analytical challenge, but rather also an opportunity to maximise complementary information from different observational approaches. It is therefore a priority to ensure that observational and analytical guidelines for flight height sampling allow for and perhaps encourage simultaneous monitoring with complimentary methods.

6. REFERENCES

Acácio, M., Atkinson, P.W., Silva, J.P. & Franco, A.M.A. 2022. Performance of GPS/GPRS tracking devices improves with increased fix interval and is not affected by animal deployment. *PLOS ONE* **17**: e0265541.

Ainley, D.G., Porzig, E., Zajanc, D. & Spear, L.B. 2015. Seabird flight behavior and height in response to altered wind strength and direction. *Marine Ornitholog*y **43**: 25–36.

Akeresola, R.A., Butler, A., Jones, E.L., King, R., Elvira, V., Black, J. & Robertson, G. 2024. Validating hidden Markov models for seabird behavioural inference. *Ecol. Evol.* **14**: e11116.

Anderson, K., Hancock, S., Disney, M., Gaston, K.J., 2016. Is waveform worth it? A comparison of LiDAR approaches for vegetation and landscape characterization. *Remote Sens. Ecol. Conserv.* **2**: 5–15.

Aschwanden, J., Schmidt, M., Wichmann, G., Stark, H., Peter, D., Steuri, T. & Liechti, F. 2020. Barrier effects of mountain ranges for broad-front bird migration. *J. Ornithol.* **161**: 59–71.

Aschwanden, J., Stark, H., Peter, D., Steuri, T., Schmid, B. & Liechti, F. 2018. Bird collisions at wind turbines in a mountainous area related to bird movement intensities measured by radar. *Biol. Conserv.* **220**: 228–236.

Auger-Méthé, M., Field, C., Albertsen, C. *et al.* 2016. State-space models' dirty little secrets: even simple linear

Gaussian models can have estimation problems. *Sci. Rep.* **6**: 26677.

Bailey, H., Brookes, K.L. & Thompson, P.M. 2014. Assessing environmental impacts of offshore wind farms: lessons learned and recommendations for the future. *Aquat. Biosyst.* **10**: 8.

Baker, B., Meadows, M., Ruffino, L. & Anderson, O.R. 2022. Towards better estimates of Manx Shearwater and European Storm-petrel population abundance and trends, demographic rates and at-sea distribution and behaviour. JNCC Research Report **719**, JNCC, Peterborough.

Band, B., 2000. Wind farm impacts on birds – calculating a theoretical collision risk assuming no avoiding action. NatureScot (Guidance note). Scottish Natural Heritage.

Band, B. 2012a. Using a collision risk model to assess bird collision risks for offshore windfarms.

Band, B. 2012b. Using a collision risk model to assess bird collision risks for offshore windfarms – with extended method: worked example. British Trust for Ornithology (BTO).

Band, B., Madders, M. & Whitfield, D. 2007. Developing field and analytical methods to assess avian collision risk at wind farms.

Barbraud, C. & Thiebot, J.-B. 2009. On the importance of estimating detection probabilities from at-sea surveys of flying seabirds. *J. Avian Biol.* **40**: 584–590.

Bodey, T.W., Cleasby, I.R., Bell, F., Parr, N., Schultz, A., Votier, S.C. & 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.

Boersch-Supan, P.H., Brighton, C.H., Thaxter, C.B. & Cook, A.S.C.P. 2024. Natural body size variation in seabirds provides a fundamental challenge for flight height determination by single-camera photogrammetry: a comment on Humphries *et al.* 2024. *Mar. Biol.* **171**: 122.

Bogdanova, M.I., Regan, C., Gunn, C., Newell, M., Benninghaus, E., Langlois Lopez, S., Quintin, M., Searle, K.R., Daunt, F., Wischnewski, S., Cleasby, I.R., Witcutt, E., Kinchin-smith, D., Holmes, E., Fox, D., McCluskie, A., Whyte, K., Butler, A. & Jones, E.L. 2021. 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.

Borberg, J.M., Ballance, L.T., Pitman, R.L. & Ainley, D.G. 2005. A test for bias due to seabird avoidance of ships when conducting surveys in the tropical Pacific. *Marine Ornithology* **33**: 173–179.

Borchers, D.L. & Cox, M.J. 2017. Distance Sampling Detection Functions: 2D or Not 2D? *Biometrics* **73**: 593–602.

Borkenhagen, K., Corman, A.M. & Garthe, S. 2018. Estimating flight heights of seabirds using optical rangefinders and GPS data loggers: a methodological comparison. *Mar. Biol.* **165**, 17.

Bouten, W., Baaij, E.W., Shamoun-Baranes, J. & Camphuysen, K.C.J. 2013. A flexible GPS tracking system for studying bird behaviour at multiple scales. *J. Ornithol.* **154**: 571–580.

Brabant, R., Rumes, B. & Degraer, S. 2021. Occurrence of intense bird migration events at rotor height in Belgian offshore wind farms and curtailment as possible mitigation to reduce collision risk. pp. 47-55 in *Environmental Impacts of Offshore Wind Farms in the Belgian Part of the North Sea: Attraction, Avoidance and Habitat Use at Various Spatial Scales.* Memoirs on the Marine Environment.

Brighton, C.H., Kloepper, L.N., Harding, C.D., Larkman, L., McGowan, K., Zusi, L. & Taylor, G.K., 2022. Raptors avoid the confusion effect by targeting fixed points in dense aerial prey aggregations. *Nat. Commun.* **13**: 4778.

Brighton, C.H., Thomas, A.L.R. & Taylor, G.K. 2017. Terminal attack trajectories of Peregrine Falcons are described by the proportional navigation guidance law of missiles. *Proc. Natl. Acad. Sci.* **114**: 13,495–13,500.

Bruderer, B., Peter, D. & Korner-Nievergelt, F. 2018. Vertical distribution of bird migration between the Baltic Sea and the Sahara. *J. Ornithol.* **159**, 315–336.

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

Bugnot, A.B., Mayer-Pinto, M., Airoldi, L., Heery, E.C., Johnston, E.L., Critchley, L.P., Strain, E.M.A., Morris, R.L., Loke, L.H.L., Bishop, M.J., Sheehan, E.V., Coleman, R.A. & Dafforn, K.A. 2021. Current and projected global extent of marine built structures. *Nat. Sustain.* **4**: 33–41.

Camphuysen, K. (C. J.), Shamoun-Baranes, J., Bouten, W. & Garthe, S. 2012. Identifying ecologically important marine areas for seabirds using behavioural information in combination with distribution patterns. *Biol. Conserv., Seabirds and Marine Protected Areas planning* **156**: 22–29.

Camphuysen, K.C.J., Fox, T.A.D., Leopold, M.M.F. & Petersen, I.K. 2004. Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the U.K. Royal Netherlands Institute for Sea Research, Texel.

Caneco, B., Humphries, G., Cook, A. & Masden, E. 2022. Estimating bird collisions at offshore windfarms with stochLAB. Software **hidef-aerial-surveying.github.io/stochLAB**

Carroll, M.J., Wakefield, E.D., Scragg, E.S., Owen, E., Pinder, S., Bolton, M., Waggitt, J.J. & Evans, P.G.H. 2019. Matches and Mismatches Between Seabird Distributions Estimated From At-Sea Surveys and Concurrent Individual-Level Tracking. *Front. Ecol. Evol.* **7**.

Certain, G. & Bretagnolle, V. 2008. Monitoring seabirds population in marine ecosystem: The use of strip-transect aerial surveys. *Remote Sens. Environ.* **112**: 3,314–3,322.

Chai, L.T., Wong, C.J., James, D., Loh, H.Y., Liew, J.J.F., Wong, W.V.C. & Phua, M.H. 2022. Vertical accuracy comparison of multi-source Digital Elevation Model (DEM) with Airborne Light Detection and Ranging (LiDAR). IOP Conf. Ser. *Earth Environ. Sci*. **1053**: 012025.

Clausen, E., Welcker, J. & Liesenjohann, T. 2023. Bioseco Pilot Study Results Report. Camera-Based Anti-Collision System. Species Protection Assessment. (No. 21_1176). BioConsult SH.

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**: 1,474–1,482.

Clements, S.J., Ballard, B.M., Eccles, G.R., Sinnott, E.A. & Weegman, M.D. 2021. Trade-offs in performance of six lightweight automated tracking devices for birds. *J. Field Ornithol.* **92**: 506–517.

Cohen, E.B., Buler, J.J., Horton, K.G., Loss, S.R., Cabrera-Cruz, S.A., Smolinsky, J.A. & Marra, P.P. 2022. Using weather radar to help minimize wind energy impacts on nocturnally migrating birds. *Conserv. Lett.* **15**: e12887.

Cole, E.-L., Waggitt, J.J., Hedenstrom, A., Piano, M., Holton, M.D., Börger, L. & Shepard, E.L.C. 2019. The Ornithodolite as a tool to quantify animal space use and habitat selection: a case study with birds diving in tidal waters. *Integr. Zool*. **14**: 4–16.

Cook, A.S.C.P., Johnston, A., Wright, L.J. & Burton, N.H.K. 2012. Strategic Ornithological Support Services Project SOSS-02: A review of flight heights and avoidance rates of birds in relation to offshore wind farms. *BTO Research Report* **618**. British Trust for Ornithology, Thetford, UK.

Cook, A.S.C.P., Ward, R.M., Hansen, W.S. & Larsen, L. 2018. Estimating seabird flight height using LiDAR. *Scott. Mar. Freshw. Sci*. **9**.

Cox, M.J., Borchers, D.L., Demer, D.A., Cutter, G.R. & Brierley, A.S. 2011. Estimating the density of Antarctic Krill (*Euphausia Superba*) from multi-beam echo-sounder observations using distance sampling methods. *J. R. Stat. Soc. Ser. C Appl. Stat.* **60**: 301–316.

Davies, A.B. & Asner, G.P. 2014. Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends Ecol. Evol*. **29**: 681–691.

Davies, J.G., Boersch-Supan, P.H., Clewley, G.D., Humphreys, E.M., O'Hanlon, N.J., Shamoun-Baranes, J., Thaxter, C.B., Weston, E. & Cook, A.S.C.P. 2024. Influence of wind on Kittiwake *Rissa tridactyla* flight and offshore wind turbine collision risk. *Mar. Biol.* **171**: 191.

de Margerie, E., Simonneau, M., Caudal, J.-P., Houdelier, C. & Lumineau, S. 2015. 3D tracking of animals in the field, using rotational stereo videography. *J. Exp. Biol.* jeb.118422.

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

Dokter, A.M., Baptist, M.J., Ens, B.J., Krijgsveld, K.L. & Van Loon, E.E. 2013. Bird radar validation in the field by time-referencing line-transect surveys. *PLOS ONE* **8**: e74129.

Dokter, A.M., Liechti, F., Stark, H., Delobbe, L., Tabary, P. & Holleman, I. 2011. Bird migration flight altitudes studied by a network of operational weather radars. *J. R. Soc. Interface* **8**: 30–43.

Domínguez del Valle, J., Cervantes Peralta, F. & Jaquero Arjona, M.I. 2020. Factors affecting carcass detection at wind farms using dogs and human searchers. *J. Appl. Ecol.* **57**: 1,926–1,935.

Donovan, C.R. & Caneco, B.A.R. 2020. Seabird survey designs for the east coast of Scotland. *Scott. Mar. Freshw. Sci.* **11**.

Duerr, A.E., Parsons, A.E., Nagy, L.R., Kuehn, M.J. & Bloom, P.H. 2023. Effectiveness of an artificial intelligence-based system to curtail wind turbines to reduce eagle collisions. *PLOS ONE* **18**: e0278754.

Dutilleux, G., Sandercock, B.K. & Kålås, J.A. 2023. Chasing the bird: 3D acoustic tracking of aerial flight displays with a minimal planar microphone array. *Bioacoustics* **32**: 622–641.

Evangelista, D.J., Ray, D.D., Raja, S.K. & Hedrick, T.L. 2017. Three-dimensional trajectories and network analyses of group behaviour within Chimney Swift flocks during approaches to the roost. *Proc. R. Soc. B Biol. Sci.* **284**: 20162602.

Evans, T.J., Young, R.C., Watson, H., Olsson, O. & Åkesson, S. 2020. Effects of back-mounted biologgers on condition, diving and flight performance in a breeding seabird. *J. Avian Biol.* **51**: e02509.

Farnsworth, A. & Russell, R.W., 2007. Monitoring flight calls of migrating birds from an oil platform in the northern Gulf of Mexico. *J. Field Ornithol.* **78**: 279–289.

Fijn, R.C. & Collier, M.C., 2022. Distribution and flight heights of Sandwich Terns *Thalasseus sandvicensis* during different behaviours near wind farms in the Netherlands. *Bird Study* **69**: 53–58.

Fijn, R.C., Krijgsveld, K.L., Poot, M.J.M. & Dirksen, S. 2015. Bird movements at rotor heights measured continuously with vertical radar at a Dutch offshore wind farm. *Ibis* **157**: 558–566.

Fischer, J.H., Bose, S., Romero, C., Charteris, M., Crowe, P., Parker, G.C., Ray, S., Rexer-huber, K., Sagar, P.M., Thompson, D.R., Bell, E., Debski, I. & Quiñones, J. 2023. Combining tracking with at-sea surveys to improve occurrence and distribution estimates of two threatened seabirds in Peru. *Bird Conserv. Int.* **33**: e41.

Fleming, C.H., Drescher-Lehman, J., Noonan, M.J., Akre, T.S.B., Brown, D.J., Cochrane, M.M., Dejid, N., DeNicola, V., DePerno, C.S., Dunlop, J.N., Gould, N.P., Harrison, A.-L., Hollins, J., Ishii, H., Kaneko, Y., Kays, R., Killen, S.S., Koeck, B., Lambertucci, S.A., LaPoint, S.D., Medici, E.P., Meyburg, B.-U., Miller, T.A., Moen, R.A., Mueller, T., Pfeiffer, T., Pike, K.N., Roulin, A., Safi, K., Séchaud, R., Scharf, A.K., Shephard, J.M., Stabach, J.A., Stein, K., Tonra, C.M., Yamazaki, K., Fagan, W.F. & Calabrese, J.M. 2020. A comprehensive framework for handling location error in animal tracking data. *bioRxiv* 2020.06.12.130195.

Forster, J., Macleod, K. & Scott, M. 2024. Response to: "Natural body size variation in seabirds provides a fundamental challenge for flight height determination by single-camera photogrammetry." *Mar. Biol.* **171**: 120.

Fox, A.D., Desholm, M., Kahlert, J., Christensen, T.K. & Krag Petersen, I. 2006. Information needs to support environmental impact assessment of the effects of European marine offshore wind farms on birds. *Ibis* **148**: 129–144.

Gayk, Z.G. & Mennill, D.J. 2020. Pinpointing the position of flying songbirds with a wireless microphone array: three-dimensional triangulation of warblers on the wing. *Bioacoustics* **29**: 375–386.

Geen, G.R., Robinson, R.A. & Baillie, S.R. 2019. Effects of tracking devices on individual birds – a review of the evidence. *J. Avian Biol.* **50**: jav.01823.

Gibb, R., Shoji, A., Fayet, A.L., Perrins, C.M., Guilford, T. & Freeman, R. 2017. Remotely sensed wind speed predicts soaring behaviour in a wide-ranging pelagic seabird. *J. R. Soc. Interface* **14**: 20170262.

Gillies, N., Fayet, A.L., Padget, O., Syposz, M., Wynn, J., Bond, S., Evry, J., Kirk, H., Shoji, A., Dean, B., Freeman, R. & Guilford, T. 2020. Short-term behavioural impact contrasts with long-term fitness consequences of biologging in a long-lived seabird. *Sci. Rep.* **10**: 15056.

Gissi, E., Manea, E., Mazaris, A.D., Fraschetti, S., Almpanidou, V., Bevilacqua, S., Coll, M., Guarnieri, G., Lloret-Lloret, E., Pascual, M., Petza, D., Rilov, G., Schonwald, M., Stelzenmüller, V. & Katsanevakis, S. 2021. A review of the combined effects of climate change and other local human stressors on the marine environment. *Sci. Total Environ*. **755**: 142564.

Glennie, R., Buckland, S.T., Langrock, R., Gerrodette, T., Ballance, L.T., Chivers, S.J. & Scott, M.D. 2021. Incorporating animal movement into distance sampling. *J. Am. Stat. Assoc.* **116**: 107–115.

Glennie, R., Buckland, S.T. & Thomas, L. 2015. The effect of animal movement on line transect estimates of abundance. *PLOS ONE* **10**: e0121333.

Grant, M.L., Bond, A.L. & Lavers, J.L. 2022. The influence of seabirds on their breeding, roosting and nesting grounds: a systematic review and meta-analysis. *J. Anim. Ecol.* **91**: 1,266–1,289.

Guo, Q., Su, Y., Hu, T., Guan, H., Jin, S., Zhang, J., Zhao, X., Xu, K., Wei, D., Kelly, M. & Coops, N.C. 2021. Lidar boosts 3D ecological observations and modelings: a review and perspective. *IEEE Geosci. Remote Sens. Mag.* **9**: 232–257.

GWEC. 2023. Global offshire wind report 2023. Global Wind Energy Council, Brussels.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R. & Watson, R. 2008. A global map of human impact on marine ecosystems. *Science* **319**: 948–952.

Harel, R., Horvitz, N. & Nathan, R. 2016. Adult vultures outperform juveniles in challenging thermal soaring conditions. *Sci. Rep.* **6**: 27865.

Harwood, A.J.P., Perrow, M.R. & Berridge, R.J. 2018. Use of an optical rangefinder to assess the reliability of seabird flight heights from boat-based surveyors: implications for collision risk at offshore wind farms. *J. Field Ornithol*. **89**: 372–383.

Hilgerloh, G. 2023. Flight altitudes of Arctic and Nordic geese in their wintering area – a radar study. *Goose Bulletin* **28**: 9–26.

Humphries, G.R.W., Fail, T., Watson, M., Houghton, W., Peters-Grundy, R., Scott, M., Thomson, R., Keogan, K. & Webb, A. 2023. Aerial photogrammetry of seabirds from digital aerial video images using relative change in size to estimate flight height. *Mar. Biol.* **170**: 18.

Hüppop, O., Ciach, M., Diehl, R., Reynolds, D.R., Stepanian, P.M. & Menz, M.H.M. 2019. Perspectives and challenges for the use of radar in biological conservation. *Ecography* **42**: 912–930.

Hüppop, O., Dierschke, J., Exo, K.-M., Fredrich, E. & Hill, R. 2006. Bird migration studies and potential collision risk with offshore wind turbines. *Ibis* **148**: 90–109.

Huso, M., Dalthorp, D., Miller, T. & Bruns, D. 2016. Wind energy development: methods to assess bird and bat fatality rates post-construction. *Human–Wildlife Interact*. **10**: 36fe-0296

Hyrenbach, K.D., Henry, M.F., Morgan, K.H., Welch, D.W., Sydeman, W.J., 2007. Optimizing the width of strip transects for seabird surveys from vessels of opportunity. *Marine Ornithology* **35**: 29–37.

Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., Jarvis, S., Lahoz-Monfort, J., Pagel, J., Pescott, O.L., Schmucki, R., Simmonds, E.G. & O'Hara, R.B. 2020. Data integration for large-scale models of species distributions. *Trends Ecol. Evol.* **35**: 56–67.

Jarrett, D., Calladine, J., Cook, A.S.C.P., Upton, A., Williams, J., Williams, S., Wilson, J.M., Wilson, M.W., Woodward, I. & Humphreys, E.M. 2021. Behavioural responses of non-breeding waterbirds to marine traffic in the near-shore environment. *Bird Study* **68**: 443–454.

Johnston, A., Cook, A.S.C.P., Wright, L.J., Humphreys, E.M. & Burton, N.H.K. 2014. Modelling flight heights of marine birds to more accurately assess collision risk with offshore wind turbines. *J. Appl. Ecol.* **51**: 31–41.

Johnston, A. & Cook, S.C.P. 2016. How high do birds fly? Development of methods and analysis of digital aerial data of seabird flight heights. *BTO Research Report* **676**. British Trust for Ornithology, Thetford, UK.

Johnston, D.T., Thaxter, C.B., Boersch-Supan, P.H., Davies, J.G., Clewley, G.D., Green, R.M.W., Shamoun-Baranes, Judy., Cook, A.S.C.P., Burton, N.H.K. & Humphreys, E.M. 2023. Flight heights obtained from GPS versus altimeters influence estimates of collision risk with offshore wind turbines in Lesser Black-backed Gulls *Larus fuscus*. *Mov. Ecol*. **11**: 66.

Kahlert, J., Leito, A., Laubek, B., Luigujõe, L., Kuresoo, A., Aaen, K. & Luud, A. 2012. Factors affecting the flight altitude of migrating waterbirds in Western Estonia. *Ornis Fenn*. **89**: 241–253.

Kim, M., Stoker, J., Irwin, J., Danielson, J. & Park, S. 2022. Absolute accuracy assessment of lidar point cloud using amorphous objects. *Remote Sens.* **14**: 4,767.

Kranstauber, B., Bouten, W., Leijnse, H., Wijers, B.-C., Verlinden, L., Shamoun-Baranes, J. & Dokter, A.M. 2020. High-resolution spatial distribution of bird movements estimated from a weather radar network. *Remote Sens.* **12**: 635.

Kreutzfeldt, J., Floeter, C., Lingner, T., Schmitz-Beuting, L., Reich, M. & Kunz, V.D. 2020. Analytical volume model for optimized spatial radar bat detection in onshore wind parks. *PLOS ONE* **15**: e0239911.

Kumagai, A., Kazama, K., Mikami, K. & Watanuki, Y. 2023. Black-tailed Gulls alter their flight height and airspeed according to wind conditions during their coastal commuting trips. *Mar. Ecol. Prog. Ser.* **723**: 201–212.

Lane, J.V., Jeavons, R., Deakin, Z., Sherley, R.B., Pollock, C.J., Wanless, R.J. & Hamer, K.C. 2020. Vulnerability of Northern Gannets to offshore wind farms; seasonal and sex-specific collision risk and demographic consequences. *Mar. Environ. Res*. **162**: 105196

Langlois Lopez, S., Clewley, G.D., Johnston, D.T., Daunt, F., Wilson, J.M., O'Hanlon, N.J. & Masden, E. 2023. Reduced breeding success in Great Black-backed Gulls (*Larus marinus*) due to harness-mounted GPS device. *Ibis* **166**: 69–81.

Largey, N. 2020. Investigating bird flight characteristics and habitat use to better understand ornithological impacts of wind farms. PhD Thesis. University of Highlands and Islands.

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

Larsen, J.K. & 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, B.G., Taylor, P.R., Miller, M.G.R., Dias, M.P., Oppel, S., Torres, L., Hedd, A., Le Corre, M., Phillips, R.A., Shaffer, S.A., Weimerskirch, H. & Small, C. 2016. Applying global criteria to tracking data to define important areas for marine conservation. *Divers. Distrib.* **22**: 422–431.

Lato, K.A., Stepanuk, J.E.F., Heywood, E.I., Conners, M.G. & Thorne, L.H. 2022. Assessing the accuracy of altitude estimates in avian biologging devices. *PLOS ONE* **17**: e0276098

Leemans, J.J., van Bemmelen, R.S.A., Middelveld, R.P., Kraal, J., Bravo Rebolledo, E.L., Beuker, D., Kuiper, K. & Gyimesi, A. 2022. Bird fluxes, flight-and avoidance behaviour of birds in offshore wind farm Luchterduinen. Bur. Waardenburg Rep. **22–078**.

Lefsky, M.A., Cohen, W.B., Parker, G.G. & Harding, D.J. 2002. Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience* **52**: 19–30.

Liechti, F., Aschwanden, J., Blew, J., Boos, M., Brabant, R., Dokter, A.M., Kosarev, V., Lukach, M., Maruri, M., Reyniers, M., Schekler, I., Schmaljohann, H., Schmid, B., Weisshaupt, N. & Sapir, N. 2019. Cross-calibration of different radar systems for monitoring nocturnal bird migration across Europe and the Near East. *Ecography* **42**: 887–898.

Liibusk, A., Kall, T., Rikka, S., Uiboupin, R., Suursaar, Ü. & Tseng, K.-H. 2020. Validation of Copernicus Sea Level Altimetry Products in the Baltic Sea and Estonian Lakes. *Remote Sens.* **12**: 4,062.

Linder, A.C., Lyhne, H., Laubek, B., Bruhn, D. & Pertoldi, C. 2022. Quantifying Raptors' Flight Behavior to Assess Collision Risk and Avoidance Behavior to Wind Turbines. *Symmetry* **14**: 2,245.

Ling, H., McIvor, G.E., Nagy, G., MohaimenianPour, S., Vaughan, R.T., Thornton, A. & Ouellette, N.T. 2018. Simultaneous measurements of three-dimensional trajectories and wingbeat frequencies of birds in the field. *J. R. Soc. Interface* **15**: 20180653.

Longarini, A., Duriez, O., Shepard, E., Safi, K., Wikelski, M. & Scacco, M. 2023. Effect of harness design for tag attachment on the flight performance of five soaring species. *Mov. Ecol*. **11**: 39.

Mancini, F., Boersch-Supan, P.H., Robinson, R.A., Harris, M.P. & Pocock, M. 2022. An introduction to model-based data integration for biodiversity assessments. JNCC, Peterborough.

Marques, T.A., Thomas, L., Martin, S.W., Mellinger, D.K., Ward, J.A., Moretti, D.J., Harris, D. & Tyack, P.L. 2013. Estimating animal population density using passive acoustics. *Biol. Rev.* **88**: 287–309.

Masden, E.A. & Cook, A.S.C.P. 2016. Avian collision risk models for wind energy impact assessments. *Environ. Impact Assess. Rev.* **56**: 43–49.

Matthiopoulos, J., Wakefield, E., Jeglinski, J.W.E., Furness, R.W., Trinder, M., Tyler, G., Mccluskie, A., Allen, S., Braithwaite, J. & Evans, T. 2022. Integrated modelling of seabird-habitat associations from multi-platform data: A review. *J. Appl. Ecol.* **59**: 909–920.

May, R., Steinheim, Y., Kvaløy, P., Vang, R. & Hanssen, F. 2017. Performance test and verification of an off-the-shelf automated avian radar tracking system. *Ecol. Evol*. **7**: 5,930–5,938.

McCann, D.L. & Bell, P.S. 2017. Visualising the aspect-dependent radar cross section of seabirds over a tidal energy test site using a commercial marine radar system. *Int. J. Mar. Energy* **17**: 56–63.

McClure, C.J.W., Dunn, L., McCabe, J.D., Rolek, B.W., Botha, A., Virani, M.Z., Buij, R. & Katzner, T.E. 2021. Flight altitudes of raptors in southern Africa highlight vulnerability of threatened species to wind turbines. *Front. Ecol. Evol.* **9**: 667384.

McGregor, R.M., King, S., Donovan, C.R., Caneco, B. & Webb, A. 2018. A stochastic collision risk model for seabirds in flight. Marine Scotland Science.

Mendel, B., Kotzerka, J., Sommerfeld, J., Schwemmer, H., Sonntag, N. & Garthe, S. 2014. Effects of the alpha ventus offshore test site on distribution patterns, behaviour and flight heights of seabirds. pp 95–110 in: *Ecological Research at the Offshore Windfarm Alpha Ventus: Challenges, Results and Perspectives*.

Mendel, B., Schwemmer, P., Peschko, V., Müller, S., Schwemmer, H., Mercker, M. & Garthe, S. 2019. Operational offshore wind farms and associated ship traffic cause profound changes in distribution patterns of Loons (*Gavia* spp.). *J. Environ. Manage.* **231**: 429–438.

Mitchell, I., Daunt, F., Frederiksen, M. & Wade, K. 2020. Impacts of climate change on seabirds, relevant to the coastal and marine environment around the UK. *MCCIP Sci. Rev*.

Mosbech, A., Johansen, K.L., Davidson, T.A., Appelt, M., Grønnow, B., Cuyler, C., Lyngs, P. & Flora, J. 2018. On the crucial importance of a small bird: the ecosystem services of the Little Auk (*Alle alle*) population in northwest Greenland in a long-term perspective. *Ambio* **47**: 226–243.

Nilsson, C., Dokter, A.M., Schmid, B., Scacco, M., Verlinden, L., Bäckman, J., Haase, G., Dell'Omo, G., Chapman, J.W., Leijnse, H. & Liechti, F. 2018. Field validation of radar systems for monitoring bird migration. *J. Appl. Ecol.* **55**: 2,552–2,564.

NIRAS. 2018. Seabird flight height field trial: using airborne LiDAR to measure the flight height of seabirds species at HOW03. Ørsted Hornsea Project Three.

ORJIP. 2022. Review of seabird monitoring technologies for offshore wind farms. The Offshore Renewables Joint Industry Programme.

Pacifici, K., Reich, B.J., Miller, D.A., Gardner, B., Stauffer, G., Singh, S., McKerrow, A. & Collazo, J.A. 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology* **98**: 840–850.

Pavón-Jordán, D., Stokke, B.G., Åström, J., Bevanger, K., Hamre, Ø., Torsæter, E. & May, R. 2020. Do birds respond to spiral markers on overhead wires of a high-voltage power line? Insights from a dedicated avian radar. *Glob. Ecol. Conserv.* **24**: e01363.

Peng, D., Hill, E.M., Meltzner, A.J. & Switzer, A.D. 2019. Tide gauge records show that the 18.61-year nodal tidal cycle can change high water levels by up to 30 cm. *J. Geophys. Res. Oceans* **124**: 736–749.

Pennycuick, C.J. 1982. The ornithodolite: an instrument for collecting large samples of bird speed measurements. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **300**: 61–73.

Pennycuick, C.J., Åkesson, S. & Hedenström, A. 2013. Air speeds of migrating birds observed by ornithodolite and compared with predictions from flight theory. *J. R. Soc. Interface* **10**: 20130419.

Pérez-Granados, C. & Traba, J. 2021. Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. *Ibis* **163**: 765–783.

Péron, G., Fleming, C.H., Duriez, O., Fluhr, J., Itty, C., Lambertucci, S., Safi, K., Shepard, E.L.C. & Calabrese, J.M. 2017. The energy landscape predicts flight height and wind turbine collision hazard in three species of large soaring raptor. *J. Appl. Ecol.* **54**: 1,895–1,906.

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

Perrow, M.R., Harwood, A., Berridge, R. & Skeate, E., 2017. The foraging ecology of Sandwich Terns in north Norfolk. *Br. Birds* **110**: 257–277.

Perrow, M.R., Skeate, E.R. & Gilroy, J.J. 2011. Visual tracking from a rigid-hulled inflatable boat to determine foraging movements of breeding terns. *J. Field Ornithol.* **82**: 68–79.

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. Environ. Manage*. **279**: 111509.

Phillips, A.C., Majumdar, S., Washburn, B.E., Mayer, D., Swearingin, R.M., Herricks, E.E., Guerrant, T.L., Beckerman, S.F. & Pullins, C.K. 2018. Efficacy of avian radar systems for tracking birds on the airfield of a large international airport. *Wildl. Soc. Bull.* **42**: 467–477.

Phillips, E.M., Horne, J.K., Zamon, J.E., Felis, J.J. & Adams, J. 2019. Does perspective matter? A case study comparing Eulerian and Lagrangian estimates of Common Murre (*Uria aalge*) distributions. *Ecol. Evol.* **9**: 4,805–4,819.

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**: 2,064–2,070.

Prinsloo, N.D., Postma, M., Ryan, P.G., Coetzee, M. & de Bruyn, P.J.N. 2021. Estimating bird flight height using 3-D photogrammetry. *J. Zool.* **314**: 174–186.

Pulido Mantas, T., Roveta, C., Calcinai, B., di Camillo, C.G., Gambardella, C., Gregorin, C., Coppari, M., Marrocco, T., Puce, S., Riccardi, A. & Cerrano, C. 2023. Photogrammetry, from the land to the sea and beyond: a unifying approach to study terrestrial and marine environments. *J. Mar. Sci. Eng.* **11**: 759.

Rahman, A., Farrok, O. & Haque, M.M. 2022. Environmental impact of renewable energy source based electrical power plants: Solar, wind, hydroelectric, biomass, geothermal, tidal, ocean, and osmotic. *Renew. Sustain. Energy Rev.* **161**: 112279.

Rees, S.E., Sheehan, E.V., Stewart, B.D., Clark, R., Appleby, T., Attrill, M.J., Jones, P.J.S., Johnson, D., Bradshaw, N., Pittman, S., Oates, J. & Solandt, J.-L. 2020. Emerging themes to support ambitious UK marine biodiversity conservation. *Mar. Policy* **117**: 103864.

Rhinehart, T.A., Chronister, L.M., Devlin, T. & Kitzes, J. 2020. Acoustic localization of terrestrial wildlife: current practices and future opportunities. *Ecol. Evol.* **10**: 6,794–6,818.

Robinson, J.L. & Jones, I.L. 2014. An experimental study measuring the effects of a tarsus-mounted tracking device on the behaviour of a small pursuit-diving seabird. *Behaviour* **151**: 1,799–1,826.

Rolek, B.W., Braham, M.A., Miller, T.A., Duerr, A.E., Katzner, T.E., McCabe, J.D., Dunn, L. & McClure, C.J.W. 2022. Flight characteristics forecast entry by eagles into rotor-swept zones of wind turbines. *Ibis* **164**: 968–980.

Ronconi, R.A. & Burger, A.E. 2009. Estimating seabird densities from vessel transects: distance sampling and implications for strip transects. *Aquat. Biol.* **4**: 297–309.

Ross-Smith, V.H., Thaxter, C.B., Masden, E.A., Shamoun-Baranes, J., Burton, N.H.K., 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. *J. Appl. Ecol.* **53**: 1,676–1,685.

Rothery, P., Newton, I. & Little, B. 2009. Observations of seabirds at offshore wind turbines near Blyth in northeast England. *Bird Study* **56**: 1–14.

Sansom, A., Wilson, L.J., Caldow, R.W.G. & Bolton, M. 2018. Comparing marine distribution maps for seabirds during the breeding season derived from different survey and analysis methods. *PLOS ONE* **13**: e0201797.

Schaub, T., Millon, A., De Zutter, C., Buij, R., Chadœuf, J., Lee, S., Mionnet, A. & Klaassen, R.H.G. 2023. How to improve the accuracy of height data from bird tracking devices? An assessment of high-frequency GPS tracking and barometric altimetry in field conditions. *Anim. Biotelemetry* **11**: 31.

Schmid, B., Zaugg, S., Votier, S.C., Chapman, J.W., Boos, M. & Liechti, F. 2019. Size matters in quantitative radar monitoring of animal migration: estimating monitored volume from wingbeat frequency. *Ecography* **42**: 931–941.

Schwemmer, P., Mendel, B., Sonntag, N., Dierschke, V. & Garthe, S. 2011. Effects of ship traffic on seabirds in offshore waters: implications for marine conservation and spatial planning. *Ecol. Appl.* **21**: 1,851–1,860.

Schwemmer, P., Mercker, M., Haecker, K., Kruckenberg, H., Kämpfer, S., Bocher, P., Fort, J., Jiguet, F., Franks, S., Elts, J., Marja, R., Piha, M., Rousseau, P., Pederson, R., Düttmann, H., Fartmann, T. & Garthe, S. 2023. Behavioral responses to offshore windfarms during migration of a declining shorebird species revealed by GPS-telemetry. *J. Environ. Manage.* **342**: 118131.

Searle, K.R., O'Brien, S.H., Jones, E.L., Cook, A.S.C.P., Trinder, M.N., McGregor, R.M., Donovan, C., McCluskie, A., Daunt, F. & Butler, A. 2023. A framework for improving treatment of uncertainty in offshore wind assessments for protected marine birds. *ICES J. Mar. Sci.* 2023.p.fsad025.

Seward, A., Taylor, R.C., Perrow, M.R., Berridge, R.J., Bowgen, K.M., Dodd, S., Johnstone, I. & Bolton, M. 2021. Effect of GPS tagging on behaviour and marine distribution of breeding Arctic Terns *Sterna paradisaea*. *Ibis* **163**: 197–212.

Shamoun-Baranes, J., Nilsson, C., Bauer, S. & Chapman, J. 2019. Taking radar aeroecology into the 21st century. *Ecography* **42**: 847–851.

Shields, M.A., Dillon, L.J., Woolf, D.K. & Ford, A.T. 2009. Strategic priorities for assessing ecological impacts of marine renewable energy devices in the Pentland Firth (Scotland, UK). *Mar. Policy* **33**: 635–642.

Signa, G., Mazzola, A. & Vizzini, S. 2021. Seabird influence on ecological processes in coastal marine ecosystems: An overlooked role? A critical review. *Estuar. Coast. Shelf Sci.* **250**: 107164.

Skov, H., Heinänen, S., Norman, T., Ward, R., Méndez-Roldán, S. & Ellis, I. 2018. ORJIP Bird Collision and Avoidance Study. Final report – April 2018. The Carbon Trust, United Kingdom.

Soanes, L.M., Bright, J.A., Bolton, M., Millett, J., Mukhida, F. & Green, J.A. 2015. Foraging behaviour of Brown Boobies *Sula leucogaster* in Anguilla, Lesser Antilles: Preliminary identification of at-sea distribution using a time-in-area approach. *Bird Conserv. Int.* **25**: 87–96.

Spear, L.B., Ainley, D.G., Hardesty, B.D., Howell, S.N.G. & Webb, S.W. 2004. Reducing biases affecting at-sea surveys of seabirds: use of multiple observer teams. *Marine Ornithology* **32**: 147–157.

Srinivasan, M.V., Vo, H.D. & Schiffner, I. 2022. 3D reconstruction of bird flight trajectories using a single video camera. *PLOS ONE* **17**: e0271618.

Stepanian, P.M., Horton, K.G., Hille, D.C., Wainwright, C.E., Chilson, P.B. & Kelly, J.F. 2016. Extending bioacoustic monitoring of birds aloft through flight call localization with a three-dimensional microphone array. *Ecol. Evol.* **6**: 7,039–7,046.

Stillman, R.A., Railsback, S.F., Giske, J., Berger, U. & Grimm, V. 2015. Making predictions in a changing world: the benefits of individual-based ecology. *BioScience* **65**: 140–150.

Suryan, R., Albertani, R. & Polagye, B. 2016. A synchronized sensor array for remote monitoring of avian and bat interactions with offshore renewable energy facilities (No. DOE-OSU-EE0005363). Oregon State Univ., Corvallis, United States.

Tamisiea, M.E., Hughes, C.W., Williams, S.D.P. & Bingley, R.M. 2014. Sea level: measuring the bounding surfaces of the ocean. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **372**: 20130336.

Tarroux, A., Weimerskirch, H., Wang, S., Bromwich, D., Cherel, Y., Kato, A., Ropert-Coudert, Y., Varpe, Ø., Yoccoz, N. & Descamps, S. 2016. Flexible flight response to challenging wind conditions in a commuting Antarctic seabird: Do you catch the drift? *Anim. Behav.* **113**: 99–112.

Tasker, M.L., Jones, P.H., Dixon, T. & Blake, B.F. 1984. Counting seabirds at sea from ships: a review of methods employed and a suggestion for a standardized approach. *The Auk* **101**: 567–577.

Thaxter, C. & Burton, N. 2009. High definition imagery for surveying seabirds and marine mammals: a review of recent trials and development of protocols. A British Trust for Ornithology Report Commissioned by Cowrie Ltd.

Thaxter, C.B., Clark, N.A., Ross-Smith, V.H., Conway, G.J., Bouten, W. & Burton, N.H.K. 2017. Sample size required to characterize area use of tracked seabirds. *J. Wildl. Manag.* **81**: 1,098–1,109.

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.K. 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.

Thaxter, C.B., Ross-Smith, V.H. & Cook, A.S.C.P. 2015. How high do birds fly? A review of current datasets and an appraisal of current methodologies for collecting flight height data: Literature review. *BTO Research Report* **666**. British Trust for Ornithology, Thetford, UK.

Tjørnløv, R.S., Armitage, M., Skov, H., Schlütter, F., Barker, M., Jørgensen, J.B., Mortensen, L.O., Thomas, K., Uhrenholdt, T. & Hass, M. 2023. Resolving key uncertainties of seabird flight and avoidance behaviours at offshore wind farms, Vattenfall. Sweden. Retrieved from **coilink.org/20.500.12592/cpz4fff**.

Urmy, S.S. & Warren, J.D. 2020. Evaluating the target-tracking performance of scanning avian radars by augmenting data with simulated echoes. *Methods Ecol. Evol.* **11**: 559–569.

Urmy, S.S. & Warren, J.D. 2017. Quantitative ornithology with a commercial marine radar: standard-target calibration, target detection and tracking, and measurement of echoes from individuals and flocks. *Methods Ecol. Evol.* **8**: 860–869.

van Bemmelen, R., Soudijn, F., Benden, D., Potiek, A., Chen, C., Hintzen, N., Wilkes, T., Van Kooten, T &, Gyimesi, A. 2021. Individual-based model lesser black-backed gulls in the Netherlands. Bureau Waardenburg, Culemborg.

van Bemmelen, R.S.A., Courtens, W., Collier, M.P. & Fijn, R.C. 2022. Sandwich Terns in the Netherlands in 2019-2021. Distribution, behaviour, survival and diet in light of (future) offshore wind farms. Bureau Waardenburg, Culemborg.

van Erp, J., Sage, E., Bouten, W., Van Loon, E., Camphuysen, K. & Shamoun-Baranes, J., 2023. Thermal soaring over the North Sea and implications for wind farm interactions. *Mar. Ecol. Prog. Ser.* **723**: 185–200.

van Erp, J.A., van Loon, E.E., De Groeve, J., Bradarić, M. & Shamoun-Baranes, J. 2024. A framework for post-processing bird tracks from automated tracking radar systems. *Methods Ecol. Evol.* **15**: 130–143.

Varbla, S., Ellmann, A. & Delpeche-Ellmann, N. 2021. Applications of airborne laser scanning for determining marine geoid and surface waves properties. *Eur. J. Remote Sens.* **54**: 558–568.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S. & Clawges, R.M. 2008. Lidar: shedding new light on habitat characterization and modeling. *Front. Ecol. Environ.* **6**: 90–98.

Wang, Z. & Menenti, M. 2021. Challenges and opportunities in lidar remote sensing. *Front. Remote Sens.* **2**.

Warwick-Evans, V., Atkinson, P.W., Walkington, I. & Green, J.A. 2018. Predicting the impacts of wind farms on seabirds: an individual-based model. *J. Appl. Ecol.* **55**: 503–515.

Watanuki, Y., Suryan, R.M., Sasaki, H., Yamamoto, T., Hazen, E.L., Renner, M., Santora, J.A., O'Hara, P.D. & Sydeman, W.J. 2016. Spatial ecology of marine top predators in the North Pacific: tools for integrating across datasets and identifying high use areas. North Pacific Marine Science Organization (PICES).

Weisshaupt, N., Hervo, M. & Haest, B. 2023. Comparison of bird migration in a radar wind profiler and a dedicated bird radar. *Remote Sens. Ecol. Conserv.* **9**: 820–828.

Weisshaupt, N., Lehmann, V., Arizaga, J. & Maruri, M. 2017. Radar wind profilers and avian migration: a qualitative and quantitative assessment verified by thermal imaging and moon watching. *Methods Ecol. Evol.* **8**: 1,133–1,145.

Weisshaupt, N., Lehtiniemi, T. & Koistinen, J. 2021. Combining citizen science and weather radar data to study large-scale bird movements. *Ibis* **163**: 728–736.

Wicikowski, A., Zein, B., Goddard, B., Sweeney, S., Coppack, T. & Mcgovern, S. 2022. Collection of seabird flight height data at an operational offshore wind farm using aircraft mounted LiDAR final report. APEM.

7. APPENDIX

7.1. Method-specific considerations for best practice recommendations

Visual surveys
Visual surveys are most suitable for monitoring areas where flight height distributions are needed for a range of species. Such surveys can provide additional information (e.g. species composition, behavioural activity) that can be used to prioritise future work. Observers are, however, just as likely to assign birds to the correct height band as to the incorrect band (Harwood *et al.* 2018, Perrow *et al.* 2017). The incorporation of rangefinders ought to improve visual surveys (e.g. training, calibration, accuracy assessment) and such devices should replace visual flight height assessments (i.e. height bands) where suitable. For example, where they don't compromise other aspects of environmental site assessment methods. Environmental and behavioural data should be simultaneously collected by observers. Although we recognise that satisfactory methods for heterogeneous detection in both vertical and horizontal dimensions are lacking, efforts should be made to employ experimental designs/survey protocols (e.g. calibration, data collection, validation) that facilitate the quantification of three-dimensional, species-specific detection probabilities, particularly by recording flight height, horizontal distance and bearing (true) from the observer.

Rangefinders
As with visual surveys, rangefinders are suitable for collecting flight height data for multiple species simultaneously and the presence of observers can help prioritise future work. Rangefinders should replace or complement visual surveys where possible or be routinely incorporated into observer training (e.g. calibration, accuracy assessment). Rangefinder data are negatively biased against low and very high flight heights (e.g. sampling limitations) and observers should follow additional methods (e.g. multiple measurements) when measuring such birds. The height of low flying (< 5 m) birds can be visually estimated to reduce the bias (Borkenhagen *et al.* 2018, Harwood *et al.* 2018). As with visual surveys, flight height data from rangefinders should be combined with behavioural observations and information to estimate species-specific detection probabilities should be obtained routinely, at minimum by recording elevation angle/flight height, azimuth/bearing, and distance to target/horizontal distance, acknowledging again that further statistical developments are needed to formally assess heterogeneous detection probability in the context of non-uniform flight height distributions. New or untested rangefinder equipment should be subject to particularly robust validation assessments to describe measurement characteristics (e.g. accuracy, precision) and identify sampling limitations (e.g. effective range) or considerations (e.g. optimal platform). Where model-specific (e.g.

Vectronix) best-practice (e.g. mounting, calibration, interference, data recording/downloading) guidance is available, it should be followed and reported. For example, handheld units need to be held perpendicular to the horizon to ensure clinometers work properly and the eye height of surveyors relative to sea surface must be record (i.e. sitting or standing).

Single-camera photogrammetry (aerial imagery/stationary platforms)
Currently proposed methods for estimating flight height via single-camera photogrammetry are not recommended due to measurement errors resulting from large natural intra-specific variation in seabird body size (Boersch-Supan *et al*. 2024). Flight height estimation from digital aerial surveys can produce substantially biased estimates of flight heights and collision risks, as well as their associated uncertainties given large measurement and sampling uncertainties. Nevertheless, single-camera methods are in principle attractive because they are economical and future developments of both sensor technology and analytical procedures may improve precision and accuracy, but further validation of this approach is required.

Stereophotogrammetry
Stereophotogrammetry is yet to be implemented as a tool for sampling flight height in the offshore environment. Stereo-camera approaches have great potential to provide accurate and precise estimates of flight heights. However, logistical challenges of working in offshore environments (e.g. large baseline distances are typically not feasible) and the need to make the analysis of the expected large volumes of imagery data (manual processing is unlikely to be feasible for continuous monitoring applications) mean that there is likely a cost/precision trade-off to operationalise these approaches in offshore environments. Stereophotogrammetry is yet to be more widely implemented as a tool for sampling flight height in the offshore environment and further validation needs to be conducted for devices that are robust enough to be deployed offshore and the image-analysis workflows associated with them.

Microphone array
Microphone arrays are yet to be implemented as a tool for sampling flight height in the offshore environment, and it is not well understood if the approaches used in terrestrial environments translate to the conditions encountered in coastal or offshore environments. However, we do consider microphone arrays here, as they have the potential to deliver night-time observations of migrating passerines and other bird groups, which are relatively poorly studied in the context of offshore collision risk.

Animal-borne tracking devices
Telemetry-based methods are most useful for species-specific data collection and can provide data on spatial, temporal, behavioural and individual variation. High-frequency (i.e. continuous) sampling results in the highest vertical and horizontal accuracy but complex programmes must be considered to overcome sampling biases (Péron *et al*. 2020). GPS devices provide estimates of flight altitude relative to MSL and values may need converting before use in CRMs; large differences in the vertical accuracy of flight altitude estimates may also exist between tag models. Barometric altimetry estimates flight height relative to the sea surface and thus requires additional calibration. It is therefore important to consistently test device accuracy and precision both pre- (e.g. stationary test, UAV) and post- (e.g. roosting, nesting, floating) deployment (Acácio *et al*. 2022, Lato *et al*. 2022, Schaub *et al*. 2023). The state-space model framework has a structure that naturally addresses the challenges of sampling errors in vertical space-use data (Péron *et al*. 2020). Models can be fitted directly to data with minimal processing thereby incorporating the full distribution of flight heights into analyses. The sample sizes required to make three-dimensional population level inferences are currently not understood and careful consideration is required when determining the number of devices to deploy (e.g. power analysis). Animal behaviour should be classified (e.g. Hidden Markov Models, Expectation–maximisation binary clustering) and considered within flight height distributions. The extent to which tracking devices alter animal behaviour should also be routinely assessed (Bodey *et al*. 2018, Geen *et al*. 2019). Appropriate statistical adjustments should be made for data sets containing repeated measurements from the same individual (e.g. mixed effects models) to account for temporal autocorrelation and/or individual-specific behaviours

Radar
Radar is most appropriate for collecting flight height data when species-specific distributions are not required and bird densities are high (e.g. migration periods). Radar systems must be properly calibrated to determine size-specific survey volumes and detection probabilities before accurate flight height distributions can be estimated (Schmid *et al*. 2019, Urmy and Warren, 2017). The calibration methods used and all technical information on radar parameters must be clearly reported. Radar data contain well know sampling errors and observation biases, especially in challenging environments such as open sea. The software within radar systems does not provide the level of processing sufficient to guarantee reliable data but a framework for processing radar data is available and provides clear and repeatable methods (van Erp *et al*. 2024). As with other sensor-based approaches, analysis pipelines need to be assessed for their influence on the derived flight height distributions (Urmy & Warren 2020).

LiDAR
LiDAR is particularly useful for sampling species-specific daytime flight height distributions at relatively inaccessible sites (e.g. remote offshore areas). Best practice recommendations are provided by (Cook *et al*. 2018) and are summarised here.

Aircraft should operate at altitudes > 300 m above mean sea level (minimise disturbance) and be piloted at speeds that ensure the digital camera can capture images with an overlap of 60% (assists with species identification). A resolution of 2 cm GSD and point density of > 10 m^{-2} should optimise survey coverage and facilitate the identification of most birds to species level. Careful consideration should be given to ensuring that sea conditions do not result in a high number of false positives (i.e. sea clutter) or negative observation bias via the removal of data below a threshold height. The LiDAR system must be calibrated such that the surveyed volume and associated detection probability are documented. The three-dimensional location of each bird is required to assess the distribution of birds relative to the survey volume. A minimum sample size of 100 birds from each site and/or seasons being considered offers an optimal balance between being able to fit a robust flight height distribution and a realistic level of effort for surveys. The use of formal analyses (e.g. power analyses) to determine sample size is however generally preferred. Ground control points should be surveyed during every flight and the LiDAR unit should not be moved from the aircraft after calibration (or be recalibrated accordingly). The matching of birds in LiDAR and imagery requires careful consideration, particularly groups of birds close to the sea surface. Species specific identification rates (where ID is definite, grouped where not) should be reported along with quality assurance/control of all processing steps.

Front cover: Sarah Kelman / BTO; back cover: Philip Croft /BTO

A review of existing methods to collect data on seabird flight height distributions and their use in offshore wind farm impact assessments

This document presents a review of existing methods for collecting seabird flight height data and their potential to produce flight height distributions that might be used in CRMs. The strengths, weaknesses, and limitations of different methods are identified and sources of measurement and sampling error, uncertainty and bias assessed. Best practice recommendations are provided for prominent methods and how data might be best utilised to inform stakeholders is considered.

Suggested citation: Feather, A.P., Burton, N.H.K., Johnston, D.T. & Boersch-Supan, P.H. 2025. A review of existing methods to collect data on seabird flight height distributions and their use in offshore wind farm impact assessments BTO Research Report **780**. BTO, Thetford, UK.

Offshore Wind Evidence + Change Programme | THE CROWN ESTATE | Reducing Seabird Collisions Using Evidence | NATURAL ENGLAND

ISBN 978-1-912642-77-9

9 781912 642779 >

Birds Science People