BTO RESEARCH REPORT 676

How high do birds fly? Development of methods and analysis of digital aerial data of seabird flight heights

Johnston, A. & Cook, A.S.C.P.

BTO Birds Science People

BTO Research Report Number 676

How high do birds fly?
Development of methods and analysis of digital aerial data of seabird flight heights

Alison Johnston & Aonghais, S.C.P. Cook

Report of work carried out by the British Trust for Ornithology
on behalf of Natural England and The Crown Estate

February 2016

The British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU
Registered Charity No. 216652

CONTENTS

FOREWORD

The purpose of this work was to develop a method for analysing digital aerial ornithology survey data, to derive species-specific flight heights. The focus of this work has been to develop an approach to analysis, rather than analysing a comprehensive dataset to derive generic flight height values. Although the project has been successful in developing such an approach to analysis, the BTO and Steering Group for the work emphasises that the values presented in this report are not intended to be used to inform assessments.

Any party undertaking an ornithological collision risk assessment should seek advice from the relevant regulators and statutory nature conservation bodies on appropriate flight height values and avoidance rates to use. This report does not constitute such statutory advice.

EXECUTIVE SUMMARY

1. The consideration of flight height is a key factor determining how seabirds interact with offshore wind farms. Of particular interest is the need to accurately estimate flight heights, in order to feed into Collision Risk Models (CRM).

2. Methodological advances have made it possible to estimate the flight heights of birds from digital aerial surveys, and there has been interest in developing continuous flight height distributions from these data (similar to previous distributions estimated from data collected during boat surveys).

3. The initial aims of this project were to: develop approaches to produce statistical distributions of seabird flight heights from digital aerial survey data; compare these distributions with those produced from boat survey data; assess the reasons for and implications of any differences between boat and aerial data for collision risk modelling; if appropriate, update boat-based distributions with data from digital aerial surveys. These objectives were delivered, but it was decided that due to limitations associated with data availability, it was not appropriate to update boat-based distributions with those produced here.

4. We considered three different approaches for producing statistical distributions of seabird flight heights from digital aerial survey data – a spline, as was used with the original boat-based models, a log-normal mixture model and a Gamma mixture model. We tested each approach using simulated data. We found that the Gamma mixture model performed well with simulated data, particularly when the observation error was large. There was a strong correlation between the true proportion at risk and the estimated proportion at risk from the modelled distribution. Consequently, this approach was selected for analysis of real data.

5. The Gamma mixture model was used to fit the flight height distributions of seven species - Manx shearwater, gannet, kittiwake, lesser black-backed gull, herring gull, great black-backed gull and Sandwich tern. These species had estimated flight heights from digital aerial survey data for at least 100 individuals. For most species, the fitted distributions differed from distributions previously estimated with boat survey data. Most species had a peak at the sea surface with the distribution from boat survey data, but a peak above the sea surface with the distribution from digital aerial survey data. There are many reasons why these two estimated distributions may vary, including different observation processes and data collection processes, analytical differences, site-specific differences, survey times in different seasons or times of day, behavioural patterns affected by the presence of boats or planes. We conducted a simulation that demonstrated that a combination of different observation processes and analytical approaches could lead to different estimated distributions from the same underlying simulated distribution. However, given the data available, it is not possible to assess the relative importance of other factors in relation to these factors.

6. Despite differing flight height distributions, the estimated proportion at potential collision height for the distributions derived from boat survey data and digital aerial data was similar for 5 out of 7 species. However, as the distributions from aerial survey data are based on limited data, it is not possible to conclude whether or not these results are applicable more widely.

7. There were substantial site-specific differences in the estimated flight height distribution for Kittiwake and this led to sizable differences in the estimated proportion at risk of collision at different sites. These differences may be driven by seasonal and spatial differences in flight behaviour. Data from only a small number of projects were analysed as part of this study and unfortunately, data limitations in the number of observations meant that it was not carry out

similar analyses for the other six species. This finding for kittiwake highlights the importance of better understanding site-specific differences in flight behaviour, and the need to collect more site-specific data to better inform this understanding.

8. The finer resolution of the digital aerial survey data has enabled the derivation of more precise estimated flight height distribution compared to boat-based surveys. These data collection protocols and analytical approaches outlined here will facilitate future development of generic continuous flight height distributions for use in impact assessments.

1. INTRODUCTION

Potential interactions between offshore renewables and marine wildlife are assessed through Strategic Environmental Assessment (SEA), Environmental Impact Assessment (EIA) and Habitats Regulations Assessment (HRA). Offshore wind farms have the potential to affect seabirds directly and indirectly. Of these, mortality caused by birds colliding with turbines is a key issue (Bradbury et al., 2014; Drewitt and Langston, 2006; Furness et al., 2013; Garthe and Hüppop, 2004), due to the high survival and longevity of seabirds. It is important to assess this risk and estimate the potential magnitude of the effect, as part of the information to inform a consent determination.

Seabird collision rates with offshore wind turbines are estimated using a collision risk model (CRM) (Band et al., 2007) which requires estimates of seabird flight heights. The flight height distribution is a key factor determining collision risk of seabirds with offshore wind farms (Desholm and Kahlert, 2005; Furness et al., 2013; Stumpf et al., 2011). A recently developed CRM called the "extended" Band model (Band, 2012) allows for bird flight heights to be input as a continuous frequency distribution, to generate more refined estimates of collision mortality by accounting for the variation in bird density and probability of collision with height across the risk area. Modelled flight height distributions can be presented with confidence intervals, reflecting spatial and temporal variability in the data, which can be used to inform the uncertainty surrounding final collision estimates.

Flight height information for marine birds has mostly been collected from baseline surveys for impact assessments. These surveys have predominantly estimated flight height from visual observers on boats or digital aerial surveys (Buckland et al., 2012; Camphuysen et al., 2004). Continuous flight height distributions have been calculated for multiple species using observational boat survey data (Johnston et al., 2014). However, the increased resolution of digital data has made it possible to identify species with a high degree of accuracy and to estimate their height above sea level. It is therefore now possible to use digital aerial survey data to create species-specific continuous flight height distributions. The use of species-specific flight height information is important since species differ greatly in their sensitivity to collision risk (Furness et al., 2013; Garthe and Hüppop, 2004; Johnston et al., 2014).

With the increased use of digital aerial surveys, there is a need to develop a methodology that can be used to derive continuous flight height distributions comparable to those derived from boat surveys. Initially, this work had four aims:

1. To develop statistical methods to produce flight height distributions from digital aerial survey data, and to apply these methods to a range of different species.
2. To compare the flight height distributions derived from boat and digital aerial survey data and to describe and evaluate the differences between them
3. Assess the implications for collision risk modelling of any significant differences identified between the distributions derived using the different survey methods.
4. If appropriate, update existing collations of flight height data and distributions (e.g. Johnston *et al.* 2014) with any new data derived within this study and derive appropriate statistical model to describe those revised datasets.

The data available for analysis were far more limited than initially anticipated, due to reasons of commercial confidentiality. This posed problems when making direct comparisons between distributions derived from boat and digital aerial surveys due to the confounding effects of survey platforms, site and seasonal differences in data collection, possible population changes over time and potential weather effects on flight behaviour. Consequently, this report focuses on developing a methodology with which to produce continuous flight height distributions for a number of species. We discuss potential reasons for differences in the flight height distributions derived using different survey platforms and consider the possible implications of these differences.

2. DIGITAL AERIAL FLIGHT HEIGHT DATA

In response to methodological and analytical advances (Buckland et al., 2012; Johnston et al., 2015; Thaxter and Burton, 2009), the use of digital aerial surveys to collect data on seabird abundance and distribution has increased in recent years. Planes fly in transects across the survey region and camera(s) mounted inside the aircraft looking out through a hatch in the underside of the plane collect data using either still photography or video imagery. Individual birds in these images are identified to species by trained image reviewers and experienced ornithologists.

With still photography, the flight height of each bird is estimated by trigonometry based on the size of the bird in the image and the known height of the plane. Confidence intervals are calculated by examining potential uncertainty in the size of the bird in the photograph.

With video footage, the flight height of each bird is estimated 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. Confidence intervals are calculated by bootstrapping the different pairs of frames and calculating a new mean for each bootstrapped sample.

Data were collated from two companies: APEM and HiDef, across a total of seven sites (see Appendix 1). APEM collect data as still photographs and HiDef collect data as digital video footage. For each bird a flight height is supplied as a point estimate with an associated error.

The uncertainty in the flight height estimate will vary for each bird with size, shape, behaviour and survey platform all affecting the error. The models will account for the uncertainty in each the estimate for each bird, so the model naturally incorporates the variation in error, some of which will be a result of the different survey platforms. Consequently, variation in error by survey platform is naturally accommodated in the models. Data analysis was conducted for seven species which had estimates of height for at least 100 individuals (Table 1). This threshold was chosen because models with fewer than 100 flight height estimates had a large degree of uncertainty around the estimated distributions.

As individuals in a flock are likely to be flying at similar heights, their heights are not independent and it is important to consider this non-independence from a statistical point of view. It is therefore important to define individuals in the same flock. An approximation to identifying individuals in a flock, is to consider all individuals within a single frame to be part of the same flock. With still photographs, each ‘group’ refers to a single still photograph, which at the sea surface has typical dimensions of approximately 180m x 140m. With video data, identifying individuals in the same frame is more complex, as most individuals occur in at least 6-8 frames. Therefore ‘group’ as defined here refers to the central 15% region of each frame of video. Therefore a group includes all individual birds which share their central frame. This region is approximately 125m x 7m. However, large flocks which cover more than 7m are generally assigned to the same ‘group’, even if their central frames vary slightly amongst the group. These areas were defined by the characteristics of the still and video images.

Table 1. List of species for which data included at least 100 individuals and the number of sites from which data were available. A frame defines a fixed unit of area on the sea surface and individuals within a frame may therefore be part of a flock and not independent measures of flight height. There was sufficient distance between the sites such that they could be considered independent.

Species	Number of individuals	Number of sites	Number of sites ≥ 100 individuals
Kittiwake *Rissa tridactyla*	1429	7	3
Lesser black-backed gull *Larus fuscus*	306	7	0
Great black-backed gull *Larus marinus*	281	6	0
Manx shearwater *Puffinus puffinus*	240	4	1
Gannet *Morus bassanus*	223	5	0
Herring gull *Larus argentatus*	222	7	0
Sandwich tern *Sterna sandvicensis*	205	3	1

We carried out some initial exploratory analyses on the data in order to investigate any potential correlation between estimated flight height and month of survey or distance to coast. To detect statistically significant variation in estimated height with season and distance to coast, we fitted a Gaussian GAM with a cyclic spline describing variation in the mean estimated height with month and a linear effect of distance to coast on mean height. The number of knots for the cyclic spline was set to three or half of the number of months of data, whichever was greatest. The gamma penalty for Generalized Cross Validation (GCV) was set at 1.4. Significant effects of month and distance to coast were assessed by F-tests and t-tests, respectively.

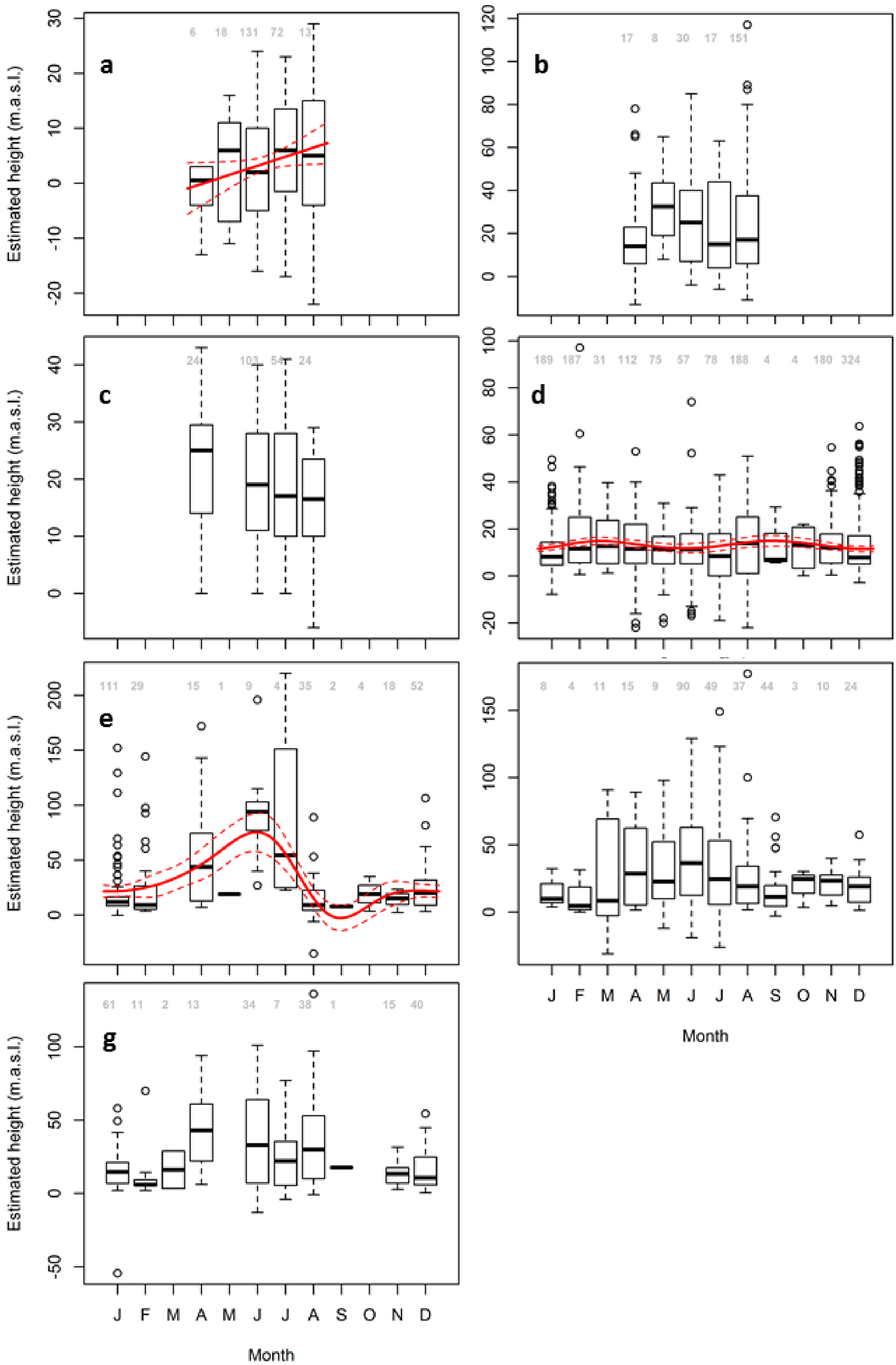

Figure 1. Exploratory analysis of the distribution of estimated height in relation to month. The numbers in grey indicate the sample size in each month. Fitted cyclic cubic splines and 95% confidence intervals (red solid and dashed lines) are shown for species which were modelled (for species with at least 10 observations in 4 months) and in which the effect of season was p<0.05. The species and significance values for the effect of season were: a) Manx shearwater p=0.042; b) Gannet p=0.717; c) Sandwich tern p=0.258; d) Kittiwake p=0.006; e) Great black-backed gull p<0.001; f) Lesser black-backed gull p=0.155; g) Herring gull p=0.713. Some values are below zero because these are the estimates of flight height from digital

aerial methods, which sometimes result in estimates below zero due to error in the estimation. These negatives are assumed to be error and the main models ensure that modelled flight heights do not occur below zero, but presented here are the raw data.

The fitted spline of month was significant for Manx shearwater, Kittiwake and Great black-backed gull (Figure 1). Manx shearwater indicated an increasing mean flight height estimate from April to September. Kittiwake had a spline that indicated very marginal effects as the magnitude of the variation was small. Great black-backed gull indicated an increasing average flight height towards the middile of the year, although there was a relatively small amount of data in the middle of the year. Flight heights are estimated with the digital aerial data with error, sometimes the point estimate for the flight height is below zero. Where this is observed, it is important to include these data in the models used to produce flight height distributions so that the full range of plausible errors are captured.

Mean flight height was significantly correlated with distance to coast for the three gull species. The magnitude of the effect was very small for all species and the precision was low. For all of these three species, estimated flight heights were slightly lower for distances further from coast. These data are not illustrated graphically due to agreement with the data owners. These exploratory analyses are reflective of the data considered in this report and should not be considered to be applicable more broadly.

3. MODELLING FRAMEWORK

3.1 Description of simulations

To assess the extent to which the statistical analysis can reproduce a true distribution, we simulated flight height distributions that were observed with error. To replicate a situation of a non-parametric distribution of flight heights (which may exist due to different behaviours in different behavioural states or locations), we used a spline to fit to a combination of 4 distributions (see Appendix 2). The fitted spline was converted to a probability distribution between 0-300m above sea level and this was assumed to be the true distribution underlying the simulation. The simulations also incorporated an observation error, which was assumed to be normally distributed around the true value. Each individual bird was assumed to have a different error distribution and each simulation had 500 individual birds (see Appendix 2 for further details of the simulations).

We initially produced a set of 100 random simulations (each with 500 birds), all with a small observation error, to test the ability of the statistical methods to infer the true flight height distribution of birds. The observation error was randomly selected for each bird and was simulated so that the average 95% confidence range was 5m. In order to understand how the magnitude of the error surrounding flight height estimates may influence the final distribution, we produced a second set of 100 random simulations, each with a large observation error, with the average 95% confidence range of 30m. Further details of the simulations are given in Appendix 2.

3.2 Methods for analysis of simulated data

There are two processes which determine estimated flight heights from digital aerial survey data.

1. The underlying flight height distribution of each species.
2. The observation process which determines how the flight height of each individual bird is estimated.

To understand the ability of the data and the models to infer the underlying distributions, we first analyse simulated data. To reflect the data, the simulations include both stage 1 (the underlying flight height distribution of the population) and stage 2 (the estimation of flight height of each bird with error). In the simulations each individual bird was independent from other birds.

To analyse the simulated data, we fitted a finite mixture model with two mixtures, and each mixture was assumed to be a Gamma distribution. A log-spline and a log-Normal mixture model were also tested, but the mixture of Gamma distributions performed best for goodness of fit. Each bird observation was determined to be drawn from one of the two mixtures (i.e. distributions):

$$mix_i \sim Categorical(p_1,\ p_2)$$

where mix_i is the mixture of individual bird *i* and this is defined by a categorical distribution with probabilities p_1 and p_2. The sum of p_1 and p_2 is constrained to be 1. Here p_1 was defined as a normal distribution on the logit scale:

$$logit(p_1) \sim N(\theta, \tau^2)$$

Each Gamma distribution $g_k(x)$ was defined by shape ω_k and rate ρ_k parameters:

$$g_k(x) \sim Gamma(\omega_k, \rho_k)$$

The mean of the distributions were constrained to be increasing with at least 0.1 between the two means, such that $\frac{\omega_1}{\rho_1} + 0.1 < \frac{\omega_2}{\rho_2}$. Weakly informative priors led to good performance of the mixture models with no evidence of label switching (Jasra et al., 2005). The priors were defined as follows:

$$logit(p_1) \sim N(0, 1.5)$$

$$\omega_k \sim Gamma(1, 0.1)$$

$$\rho_k \sim Gamma(1, 1)$$

Analyses of the 200 simulations were run in JAGS (Plummer, 2003) using the Gamma mixture model described above. Models were fitted by Markov Chain Monte Carlo (MCMC) using three chains, 250 000 iterations and a burn-in of 50 000 iterations (which was visually assessed to be a very conservative estimate of convergence time in several simulations). Automated thinning was used to reduce autocorrelation. Convergence was assessed with the BGR statistic, which (unlike several other diagnostic criteria) will identify non-convergence if there is label switching of the two mixture distributions. Credible intervals were calculated from the 2.5th and 97.5th quantiles of the posterior distributions.

To assess the models applied to the simulated data we compared the true simulated distributions to the estimated distributions in a number of ways.

1. Visual comparison of the distributions
2. Assessment of the proportion of credible intervals that contain the true values and how this varies with metres above sea level
3. Assessment of the error between estimated and true values and how this varies with metres above sea level
4. Comparison of the true proportion at potential collision height with the estimated proportion at potential collision height.

3.3 Validation of models using simulated data analysis

Convergence was reached for 80/100 of the simulations with a small observation error and 65/100 of the simulations with a larger observation error. Visual assessment of the converged distributions revealed broad agreement between the true and estimated distributions. The credible intervals for the simulations with a large observation error were larger (Figure 2). Total error between true and estimated distributions was greater for the large observation error (Figure 3), so the models were on average further from the true distributions. However, the larger credible intervals from the models with large observation error led to a greater proportion of true values within the credible intervals (Figure 4), so the models were more conservative when observation error was large. Conversely, the models produced confidence intervals that were too narrow when observation error was small. At high heights above sea level, the proportion of true values within the credible intervals was generally very low (Figure 4).

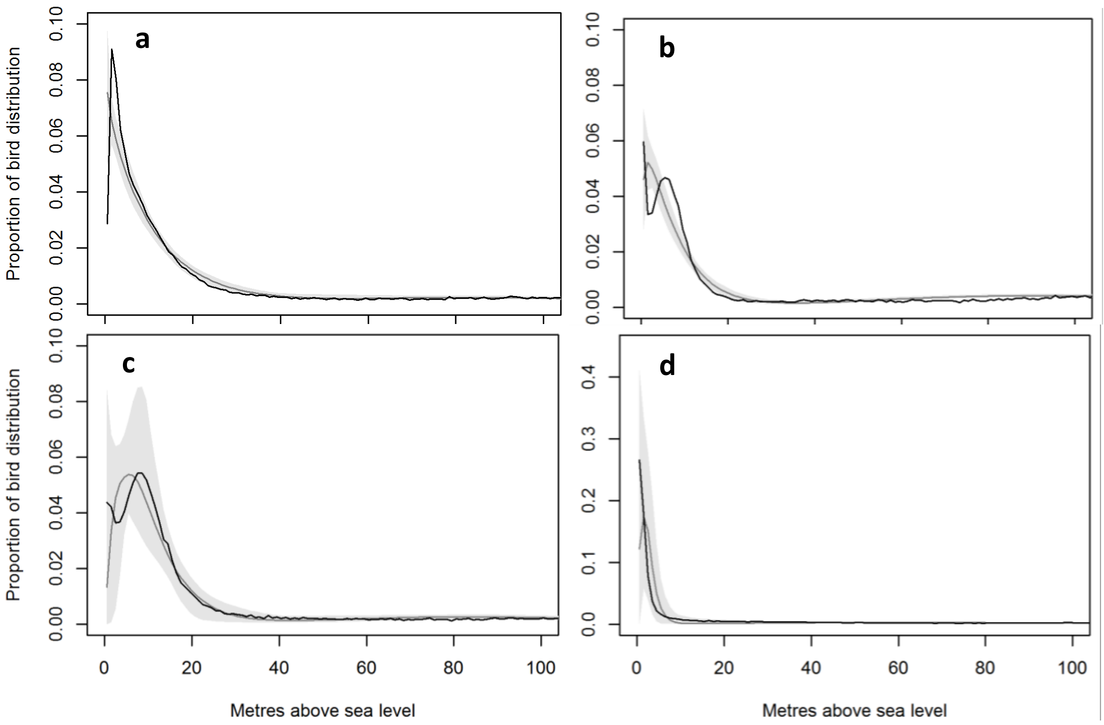

Figure 2. Examples of four simulated Spline distributions (black line) with the estimated distributions from a Gamma mixture model (grey line) and associated 95% credible intervals (pale grey polygon). Examples are for a mean standard error of the observation distribution of 5 (a, b) and 30 (c, d).

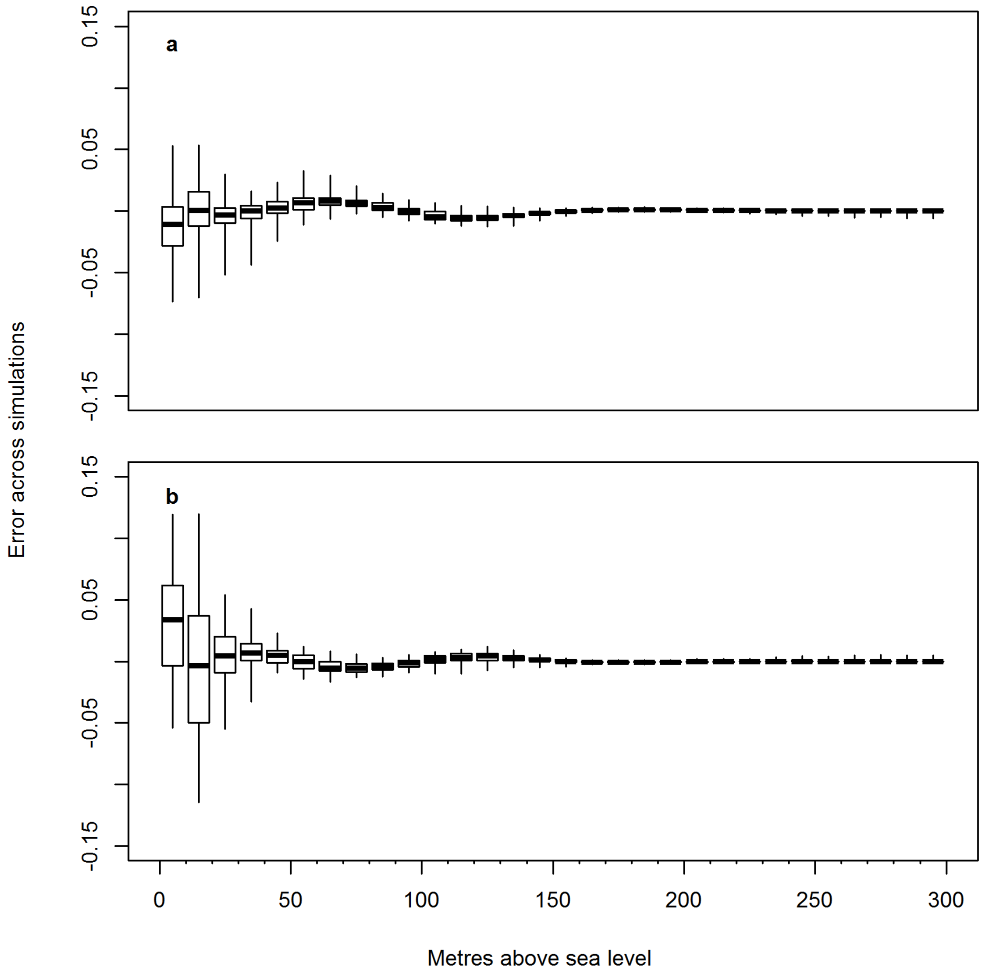

Figure 3. Total error summed into 10m categories across 100 simulations for a) small observation error (mean sd=5) and b) large observation error (mean sd=30).

To assess the sensitivity of the proportion of species at potential collision risk height, we compared the true proportion at potential collision height with the estimated proportion at potential collision height. For the simulations, potential collision height we defined as 20-120m. The correlations between the true proportion at potential collision height and the median proportion at potential collision height from the posterior distribution were r=0.98 and r=0.94 for the small and large observation error simulations, respectively (Figure 5). The 95% credible intervals around the estimated proportion at potential collision height were calculated by taking the 2.5th and 97.5th quantiles of the proportion at potential collision height from the posterior distribution. 81% and 94% of the credible intervals contained the true simulated value for the small and large observation error simulations. Although some of the credible intervals did not contain the true values (Figure 4), it is clear that the estimated proportion at potential collision height is a more robust metric to derive from these models (Figure 5).

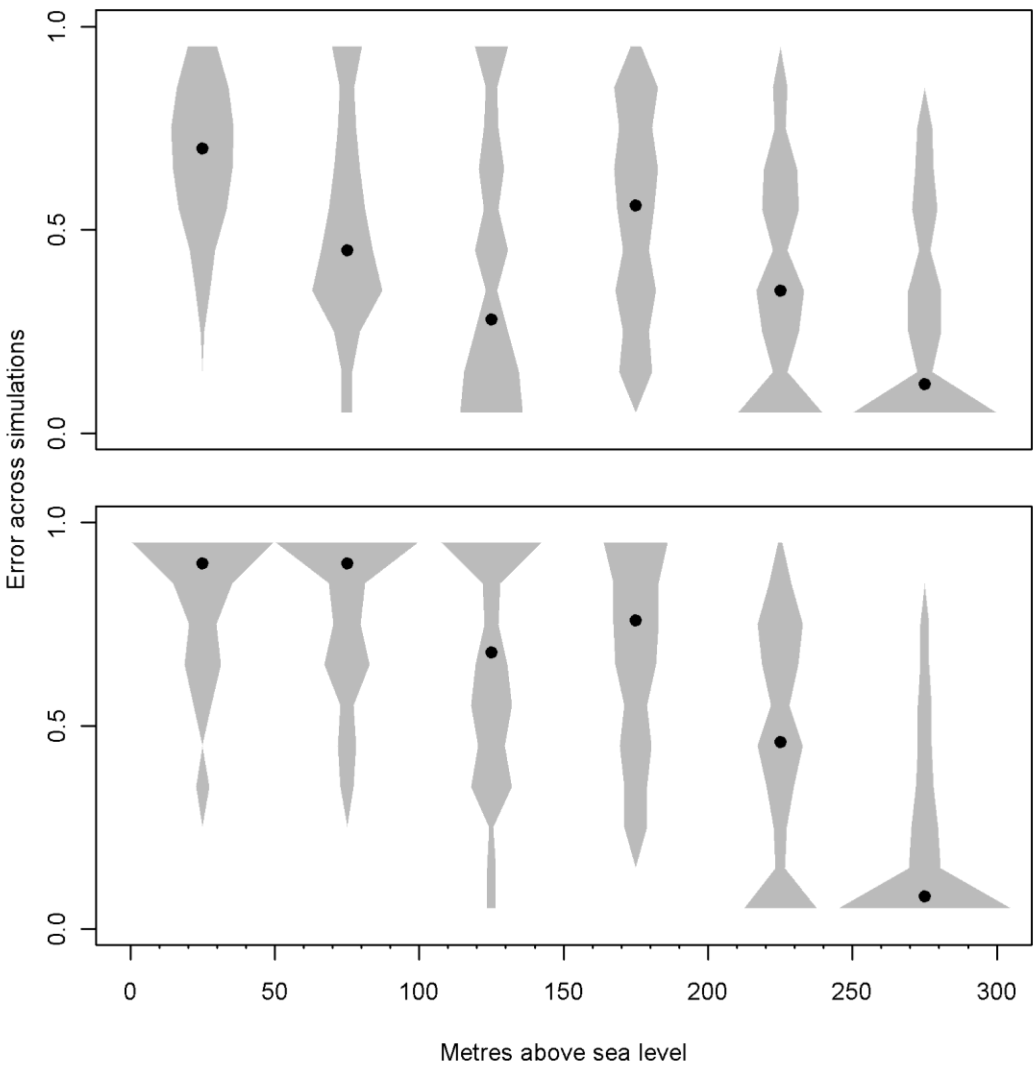

Figure 4. Leafplots of the proportion of 95% credible intervals that contain the simulated true value for: a) small observation error (mean sd=5) and b) large observation error (mean sd=30). Median values for each 50m height category are shown as black points. The grey boxes show the distribution across all simulations of the proportion of estimated confidence intervals that contain the true value. For the models with large observation error, the majority of confidence intervals contain the true value as the larger grey areas are near 1. Between 250-300m there are a low proportion of the estimated confidence intervals that contain the true value, regardless of the observation error. Generally, the confidence intervals estimated from the simulations with larger observation error (lower graph) contain a higher proportion of the true values.

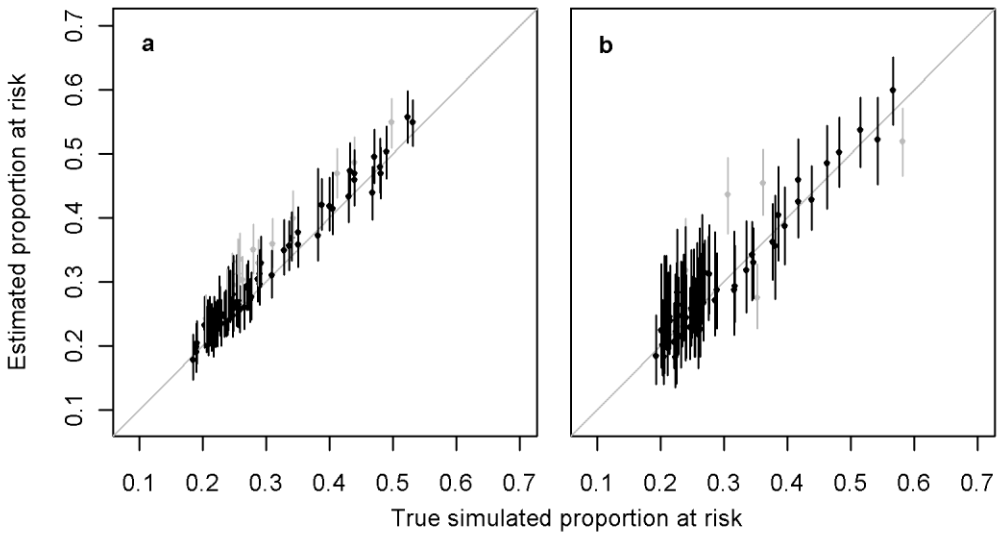

Figure 5. Comparison of true simulated proportion at potential collision height (20-120m) and Estimated proportion at potential collision height with 95% credible intervals for: a) small observation error (mean sd=5) and b) large observation error (mean sd=30). Estimated proportions and credible intervals that contain the true value are shown in black and overlap the grey line. Estimated credible intervals that contain the true value are a) 65/80 = 81% and b) 60/65 = 94%. Estimated proportions and credible intervals that do not contain the true value are shown in grey.

4. ANALYSING BIRD DATA

4.1 Methods for analysis of bird data

Data from seven species (Table 1) were analysed using the methods described in section 3 above to estimate species-specific flight height distributions. Analyses of the 200 simulations were run in JAGS (Plummer, 2003) using the Gamma mixture model code in Appendix 3. The modelling procedure followed that described in section 3.2 for the simulated data.

The flight heights of birds in groups are unlikely to be independent and therefore it would be ideal to account for this in the analysis. Birds were grouped (as described above) into those in the same frames. The distribution of number of birds in each 'group' (i.e. frame) was highly positively skewed for each species (Figure 6). We attempted several approaches to incorporate the non-independence in flight height in groups. Firstly, we attempted to include group size as a random variable in the model and group size was dependent on height. Each group was modelled with a mean flight height. The flight height of each individual bird was drawn from a distribution with a group-specific mean and constant variance. This model did not converge, which is likely due to the highly skewed distribution of group sizes and the high number of groups with only one or two individuals, which makes it challenging to estimate group variances. Secondly, we attempted to define two types of groups: large groups and small groups. The probability of a group being large or small was related to the mean flight height of the group. This model also did not converge, which is likely due to the extremely small number of 'large' groups. Treating each group as a single datapoint without considering the group size, may have resulted in estimated distributions that were not representative of the population, given that group size and flight height are likely correlated. Given the very skewed distribution of group sizes and the consequent challenges of convergence, the results below are from models in which each individual bird as was assumed to be independent.

Each estimate of bird flight height is has an associated estimate of error and these vary across individuals. The error varies with height for some species (Figure 7) and also varies slightly with the method of flight height estimation (digital stills or videos). The analytical method described above in section 3.2 incorporates the individual estimate of error in the analysis. The 95% confidence intervals for the true flight height are automatically incorporated: estimates with a narrow 95% confidence interval are assumed by the model to be nearer the true value, than estimates with wide confidence intervals. Therefore these differences in precision are fully accommodated in the modelling process. Note that in some instance the algorithms used to estimate bird flight heights may result in confidence intervals which overlap with zero (i.e. imply that birds are flying under the sea surface, see figure 7). Whilst these estimates are clearly wrong, it is important that these data are included in the analysis so that the models fully capture the plausible error associated with estimates of seabird flight heights from digital aerial survey data.

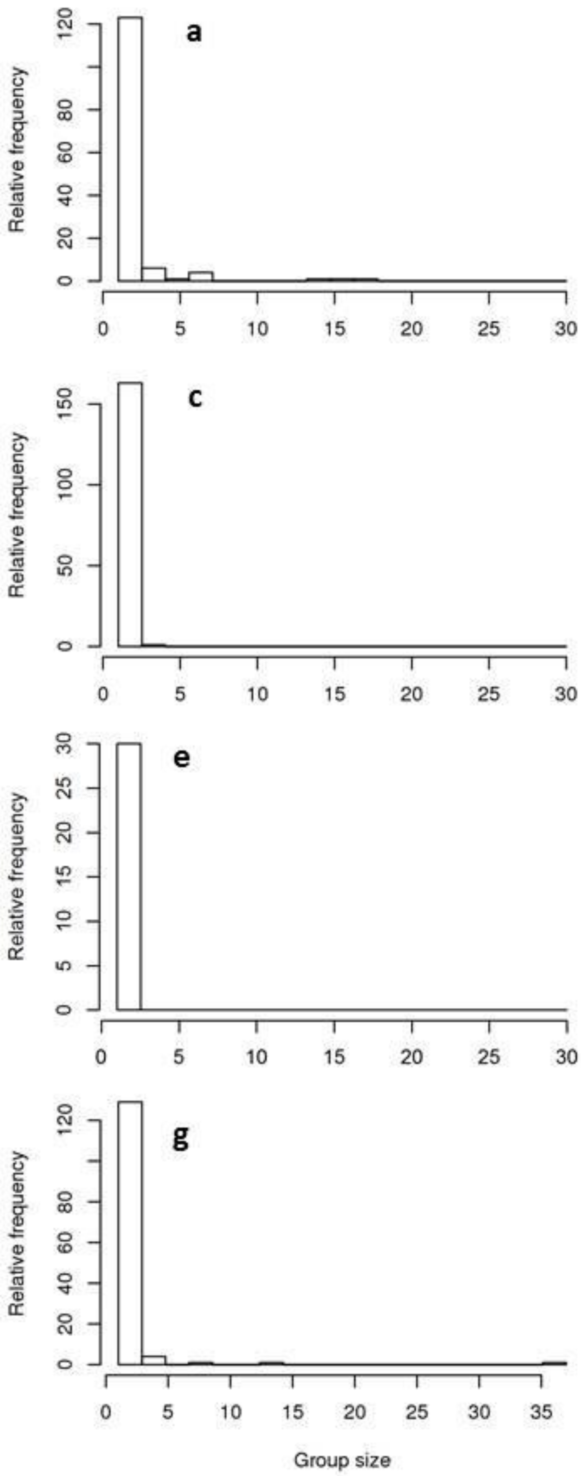

Figure 6. Distribution of group sizes (number of individual birds per frame) for a) Manx shearwater, b) Gannet, c) Sandwich tern, d) Kittiwake, e) Great black-backed gull, f) Lesser black-backed gull, g) Herring gull.

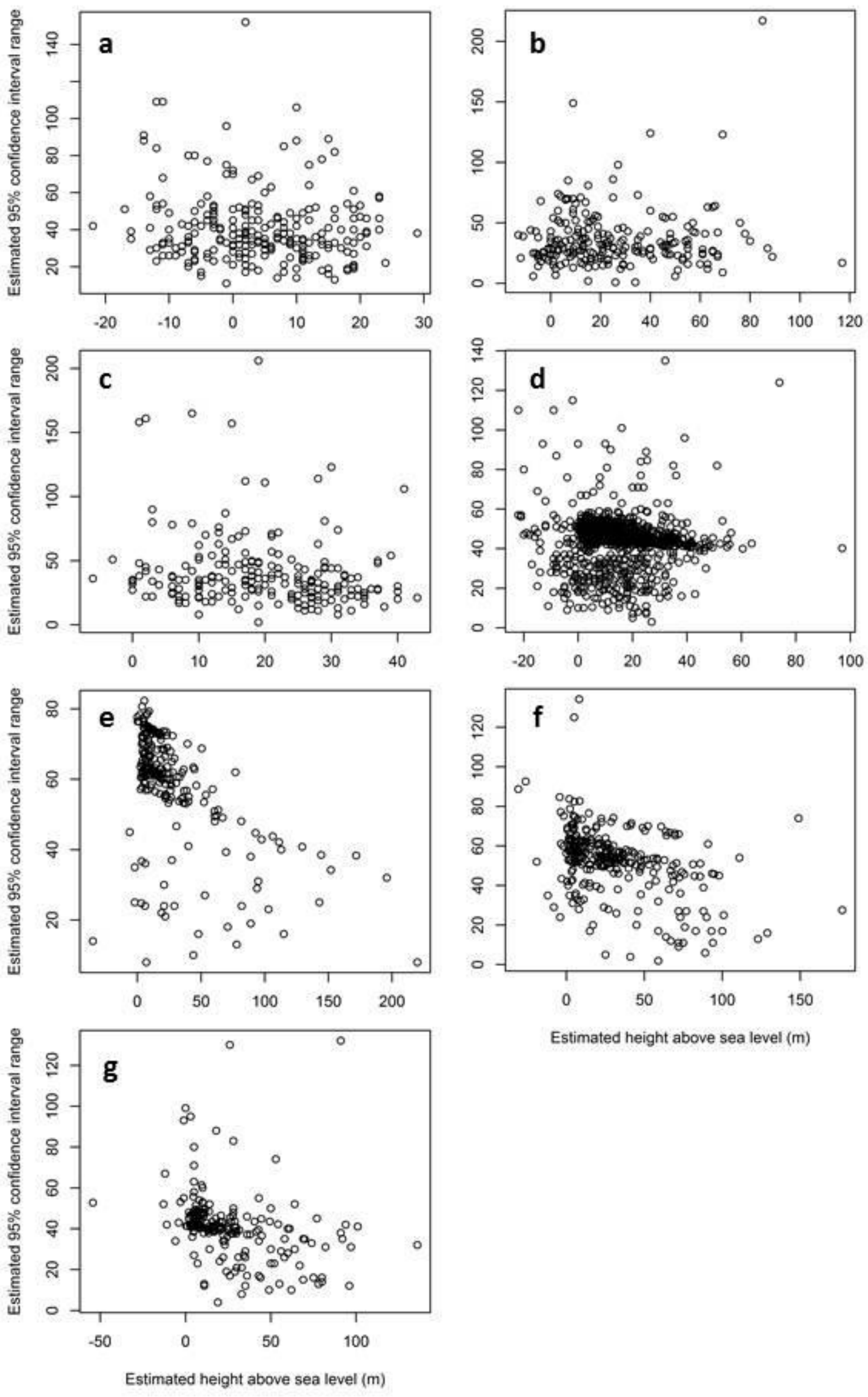

Figure 7. Relationship between estimated height above sea level and the 95% confidence interval range for a) Manx shearwater, b) Gannet, c) Sandwich tern, d) Kittiwake, e) Great black-backed gull, f) Lesser black-backed gull, g) Herring gull.

4.2 Bird flight height distributions

We plotted the estimated distributions and 95% credible intervals (Figures 8-14) for seven species (Table 1). We also show for comparison the curve that was previously estimated from boat survey data (Johnston et al., 2014). Most of the flight height distributions calculated from boat survey data have the peak of flight heights at 0m or very close to the sea surface, whereas most of the distributions from digital aerial survey data have a primary or secondary peak above the sea surface (Figures 8-14). There are two potential explanations for this. Firstly, the lower resolution of the boat survey data may mean that this secondary peak is not captured as it falls within one of flight height bands to which birds in flight are assigned. This may be particularly important for species such as Manx Shearwater which take advantage of the air currents generated by waves to minimise the energetic cost of flight (Spear & Ainley 1997). Alternatively, it may reflect a response to the survey platform. Digital aerial surveys appear to show a greater proportion of birds on the sea surface than is the case for boat surveys (WWT 2012). In the case of gulls (Figures 11-14), the presence of a peak close to the sea surface in the boat survey data may reflect the response of the birds to the survey platform as they are either flushed from the sea surface, or attracted to the boat as a potential foraging opportunity. The height of the planes used for digital aerial surveys means that they are unlikely to influence the behaviour of birds in this way (Buckland *et al.* 2012).

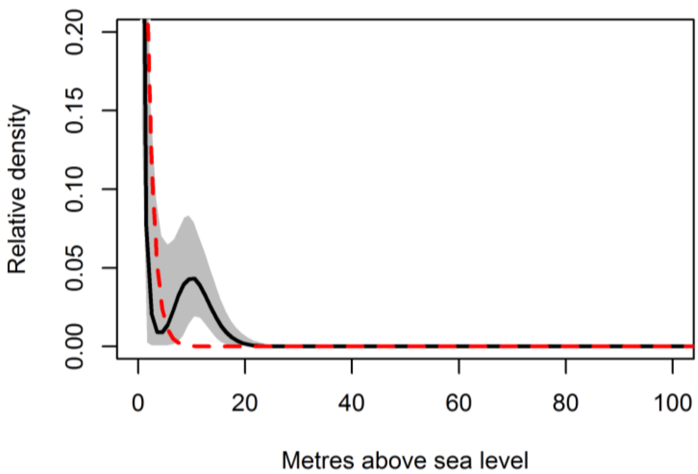

Figure 8. Estimated average flight height distribution for **Manx shearwater** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

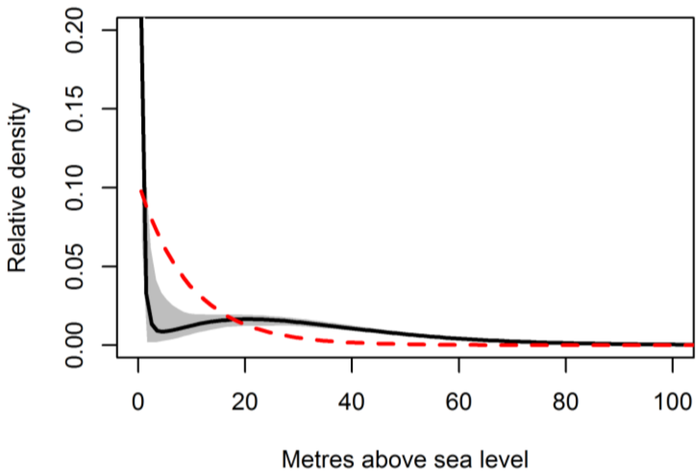

Figure 9. Estimated average flight height distribution for **Gannet** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

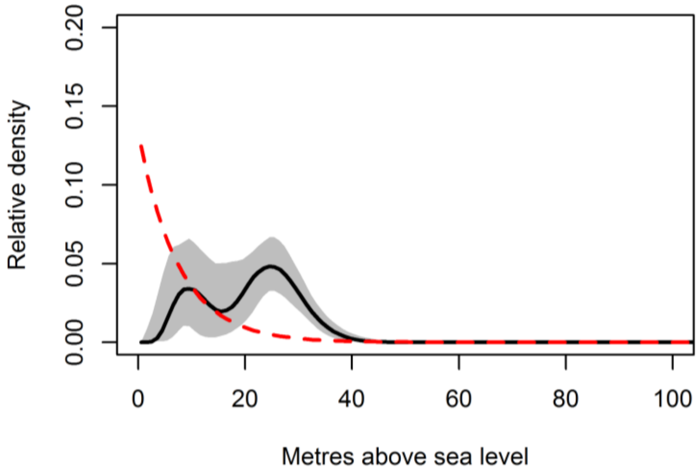

Figure 10. Estimated average flight height distribution for **Sandwich tern** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

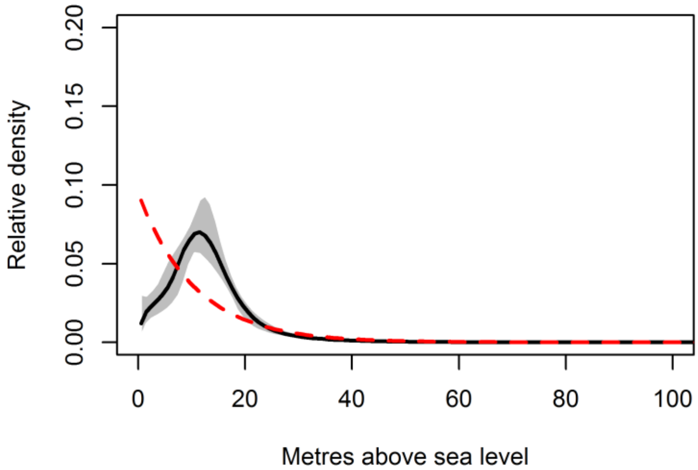

Figure 11. Estimated average flight height distribution for **Kittiwake** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

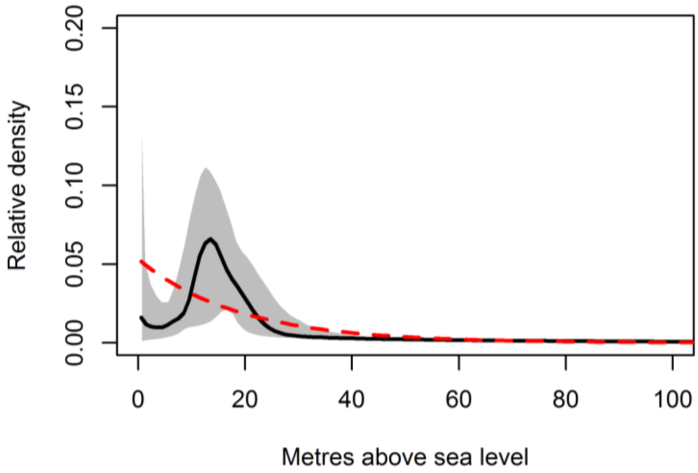

Figure 12. Estimated average flight height distribution for **Great black-backed gull** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

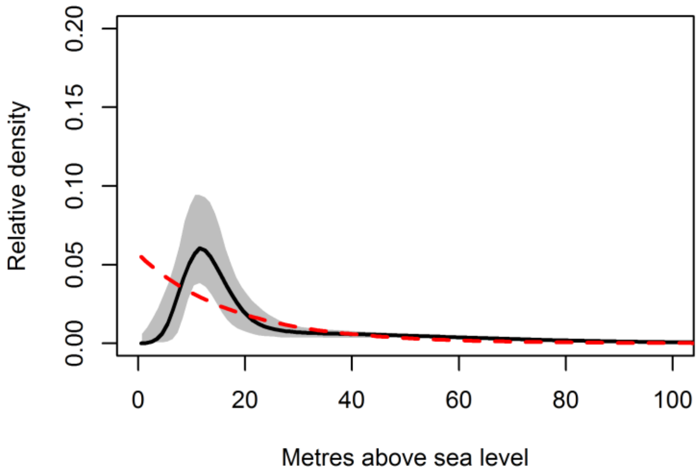

Figure 13. Estimated average flight height distribution for **Lesser black-backed gull** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

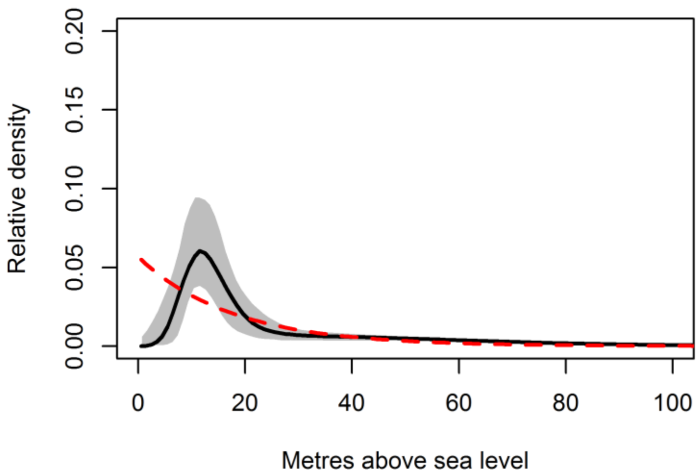

Figure 14. Estimated average flight height distribution for **Herring gull** (black line) with 95% credible intervals (grey area). The previously estimated curve from the boat survey data distributions is shown as a dashed red line.

4.3 Site-specific distributions

As at least 100 observations were required in order to derive flight height distributions from digital aerial survey data, it was only possible to derive site-specific distributions for Kittiwake from three sites (sites 1, 2 and 3 Appendix 1). Previous analyses (see section 2) suggest that the level of uncertainty surrounding distributions derived from less than 100 observations would lead to little certainty about site-specific differences. Most species did not have sufficient data to calculate site-specific flight height distributions, although Kittiwake had at least 100 observations at three sites (Table 1). To examine the extent to which there may be site-specific differences for Kittiwake, we ran the model described above for each of the three sites.

The three sites had distinctly different estimated flight height distributions (Figure 15). For the two sites with the most different distributions there was little overlap in the 95% credible intervals, suggesting clear differences between the distributions (Figure 16). These differences may relate to either site-specific differences in bird behaviour or to seasonal differences in data collection at the three sites. For example, data at site one were only collected in August whereas data from sites two and three were collected over the course of a full year.

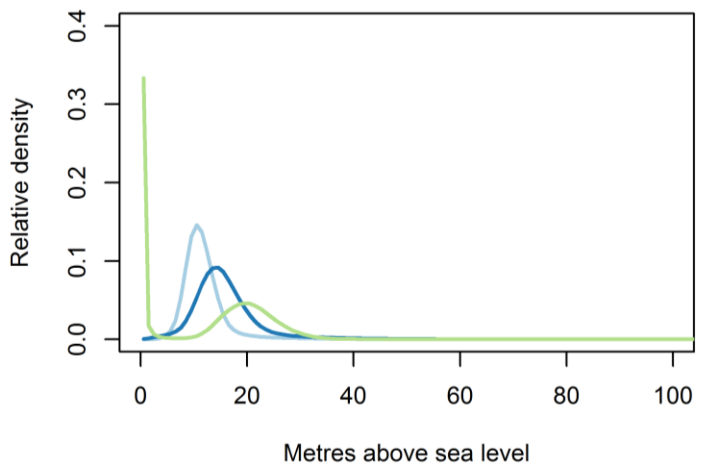

Figure 15. Estimated average flight height distribution for **Kittiwake** for the three sites (green = site 1, light blue = site 2, dark blue = site 3) with at least 100 individual birds with estimated flight heights.

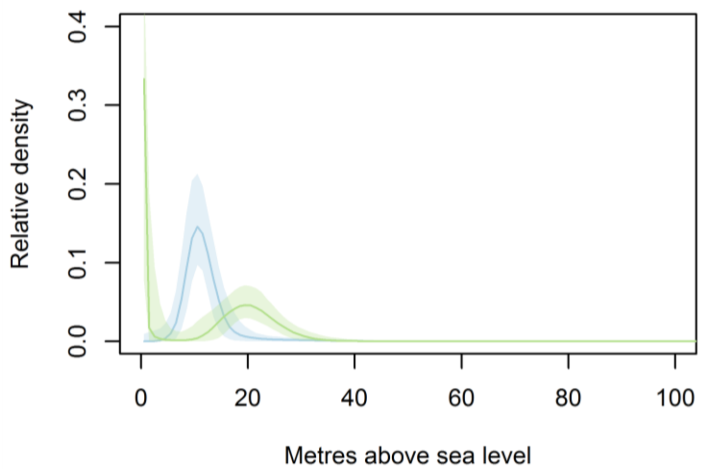

Figure 16. Estimated average flight height distribution and 95% credible intervals for **Kittiwake** for two of the three sites (green = site 1 green, light blue = site 2) with at least 100 individual birds with estimated flight heights. For most of the distributions below 30m there is no overlap in the 95% credible intervals at the two sites.

4.4 Discussion of the differences between boat and digital aerial survey data distributions

It is notable that there are a number of differences between the estimated flight height distributions from digital aerial survey data and those previously estimated from boat survey data. There are many possible reasons for these differences in estimated distributions. However, as the distributions are based on a relatively limited number of surveys and observations, which were carried out at different locations and over different time periods, it is not possible to assess the reasons for these differences statistically as any explanatory variables are likely to be confounded in a small dataset. Here we list five categories of possible reasons for this variation in estimated distributions from boat and digital aerial survey data:

1. Differences in the flight height distributions between sites
2. Differences in the flight height distributions between times or seasons
3. Differences in bird response behaviour with the different survey platforms
4. Differences in the accuracy of the different survey methods
5. Differences in the analytical methods

1. Differences in the flight height distributions between sites
In section 4.3 we highlight the differences in the estimated flight height distributions of kittiwake between sites (derived from digital aerial survey data). These differences may be due to differences in weather, season, habitat use or behaviour at the different sites. Alternatively they may be a reflection of the natural variability in flight height distributions. Given the level of uncertainty associated with distributions derived from fewer than 100 observations, such an analysis was not possible for other study species. Data contributing to the boat survey based analyses were collected from a broad range of sites (Johnston et al., 2014), whereas the digital aerial data contain greater information about a smaller range of sites. This may lead to the possibility that differences between the distributions derived from different platforms result from genuine site-specific differences in the heights at which birds fly at the various sites sampled from the two differing survey platforms. This is particularly likely to be the case if the distributions from either platform are produced with data from a small number of sites, as the individual sites will always show deviation from an average distribution.

In previous work (Johnston *et al*. 2014), in order to determine how well the modelled distributions from the boat data could be applied to novel sites a jacknife procedure was used, whereby each site in turn was dropped from the analysis and the remaining data used to estimate a flight height distribution (Johnston *et al.* 2014). The data from the excluded site was then compared to the resultant distribution in order to determine whether the proportion of birds within each flight height band fell within the 95% CIs of the modelled distribution. For the six of the seven species considered here modelled boat-based distributions were a poor match for novel sites (39-59% of independent observations within modelled 95% CIs for all species except Manx shearwater). This suggests that given there are also differences between the boat datasets collected from different sites. There may also be genuine differences between the flight height distributions at sites used for the analysis of boat data and those used for the analysis of digital aerial survey data. This would be consistent with the site-specific differences in kittiwake flight height distributions described here. These analyses suggest that for the three kittiwake sites, flight height distributions were more similar at the two sites which are closest together, and further offshore, than the distribution at the site closest to shore, although this suggestion must be treated with caution given that it is drawn from only three sites.

2. Differences in the flight height distributions between times or seasons

Data from boat surveys were collected as part of assessments of the potential impacts of proposed offshore wind farms, and will, consequently have year round coverage (Cook et al., 2012; Johnston et al., 2014). However, the times of day of collection may differ, primarily because boat surveys take a long time to cover the area of a proposed offshore wind farm (typically one to three days). A digital aerial survey is typically completed within a day and frequently in only part of a day. The result is that boat surveys typically start from first light and continue to the end of the daylight whereas a digital aerial survey will start with mobilisation from an airfield early in the day, start on-survey an hour or two later and finish with sufficient time to return to land in daylight. Collection of digital aerial survey data can also be restricted at times of day when the sun is low in the sky, leading to glare in the cameras although this effect can be managed with the angle that is flown relative to the sun and any angle placed on the camera mount. Whilst observers on boats may not be restricted in this way, sun glare may still affect their ability to detect and correctly identify birds and to correctly estimate flight height or distance. Consequently, there may be some differences in the times of day in which the data are collected and the reliability of data collected at certain times of day. These differences may be important if flight behaviour varies over the course of a day. For example, birds may time their foraging trips to coincide with tidal cycles (Irons, 1998), or have different levels of flight activity throughout the day (Daunt et al., 2002). These behavioural patterns may lead to differences in the flight heights of birds detected from boat and digital aerial surveys if one platform is better able to avoid systematic bias than the other.

Flight behaviour may also differ according to weather conditions. Health and safety concerns limit boat-based surveys to conditions where sea-state is less than 5 (Camphuysen et al., 2004). Technological limitations may similarly restrict digital aerial survey data as good visibility is required in order to detect birds. However, digital aerial surveys are typically carried out over shorter time periods than boat surveys and do not require prolonged periods of stable conditions. Consequently, digital aerial surveys may be able to capture weather conditions temporally nearer to more adverse weather conditions than boat-based surveys and therefore better reflect climatic variability. However, at present, the influence of weather conditions on seabird flight heights remains relatively uncertain.

Whilst boat survey data were collected throughout the year, digital aerial survey data used in this report tend to focus on the breeding season (Appendix 1), although two of the seven sites included in the data analysed here did include year-round data collection. Analysis of kittiwake data revealed differences in the flight heights between sites where bird abundance peaked during the breeding season and those where activity peaked over winter. However, it was not possible to separate any seasonal effects from site effects. Detecting any seasonal effects from surveys may be complicated by the fact that multiple years of data are often required to properly capture the environmental variation that may influence flight behaviour at a site (Grecian *et al.* 2010). As both boat and digital aerial survey data collection are typically limited to 2 years monitoring pre-construction, it is likely that neither method fully captures the range of flight heights used by a species at any given site.

Potential differences in seasons or times of day for the two different data collection methods may be important if birds' behaviour varies over the course of the day or the year. Behavioural patterns that alter with season or time of day may lead to a difference in the average flight heights of the data collected from boat and digital aerial surveys.

3. Differences in bird response behaviour with the different survey platforms

The height at which the aerial survey planes fly mean that birds are less likely to be attracted to, or disturbed by, the survey platform than boats (Buckland et al., 2012). There is the potential for boat-based surveys to negatively bias estimates of species flight heights as birds may fly at lower altitudes when attracted to boats (Krijgsveld et al., 2011) or be recorded by observers having just taken to flight after being flushed from the sea surface (Johnston et al., 2015; Schwemmer et al., 2011). The distributions for most species estimated from the boat survey data are positively skewed and have the highest density at zero. This is potentially consistent with negatively biased estimates of altitude from boats, as birds fly lower near boats or are flushed from the sea surface.

There are species-specific differences in the extent to which modelled estimates of flight height from digital aerial surveys differ from modelled estimates from boat-based surveys. These differences may relate to the ecology of the species concerned. Kittiwake, Herring gull, Lesser black-backed gull and Sandwich tern all have higher numbers estimated near the sea surface from the boat surveys, compared to the digital aerial surveys. Gulls in particular may be attracted to boats as they have learned to associate fishing vessels with feeding opportunities (Sotillo et al., 2014; Tyson et al., 2015). Consequently, the flight height of these birds when attracted to survey vessels may be atypical of their flight heights elsewhere in the marine environment. In contrast, a far greater proportion of gannets were near the sea surface from digital aerial surveys than from boat surveys. This finding is consistent with information presented in WWT (2012) which suggested that more gannets were recorded in flight from boat-based surveys than was the case in visual aerial surveys. The digital aerial survey data may therefore be picking up a greater proportion of birds as they taking off from the sea surface or commuting between breeding colonies and foraging areas (Cleasby *et al.* 2015), while boat-based surveys are detecting a greater proportion of birds already in flight or actively foraging. Whilst gannets may make use of the feeding opportunities provided by fishing vessels, they may be less likely to do so than gulls or kittiwakes (Camphuysen, 2011; Hudson and Furness, 1989; Krijgsveld et al., 2011). The boat surveys may be detecting birds attracted to the survey vessel by perceived feeding opportunities but the flight behaviour of these birds may not be typical of the behaviour of gannets elsewhere in the marine environment where boats (e.g. fishing boats) are absent. Better data on the at-sea flight behaviour of gannets, and the other species covered by this report, are needed in order to better understand the reasons for these differences. This might, for example, be collected using altimeters or GPS telemetry (e.g. Corman and Garthe, 2014; Cleasby *et al.* 2015) but even these techniques have limitations including the small sample numbers of birds that are fitted with such devices, the potential for such devices to affect the behaviour of the birds and their inherent measurement errors (see Thaxter *et al.* 2015).

4. Differences in the accuracy of the different survey methods

All ecological surveys measure a target parameter with some error, whether population density, annual survival or flight height. Boat surveys and digital aerial surveys have very different structures to the error and very different precision of the flight height estimate, and these may lead to different estimated distributions.

Boat surveys use human observers to estimate the flight height of birds. Although observers are often trained according to industry standards (Camphuysen *et al.,* 2004) to identify birds in offshore conditions, they are not specifically trained in height estimation, and may not detect birds that are particularly high, particularly low, or those that are further from the boat (up to 300m horizontal distance from the boat). Furthermore, they may estimate the flight height of detected birds with some bias or error, which is unlikely to be independent of the true height of the bird. There have been few studies, onshore or offshore, that have assessed the ability of observers to accurately estimate the height of birds, so it is difficult to suggest levels of bias. Pearce-Higgins et al. (2009) found there was more consistency amongst observer estimates at low heights, suggesting that height is easier to estimate low heights and more challenging at greater heights. Results from recent RSPB trials (RSPB unpublished data) suggest that observers assign birds to the correct height band between only 30%-50% of the time. In other words observers are as likely as not to assign birds to the wrong flight height band. That is a very severe limitation to the reliability of boat-based flight height data and of any modelled flight height distribution derived from such data. Further, there are likely to be differences in the skill of observers at detecting birds and estimating their height.

The estimate of flight heights within height categories is likely to reduce error and larger categories should lead to more accuracy in assigning birds to height bands. But broader height categories mean that there is lower precision and it is more challenging to model detailed variation in the flight height distribution. This coarse data format may lead to little information about detailed variation in the flight height distribution and perhaps leads to a parsimonious fitted distribution with a simple shape, similar to the exponential shape fitted for many of the species (Johnston et al., 2014).

Digital aerial surveys also have error associated with the estimate of flight height, but the structure of the error is very different to the boat survey error structure. The aim of the modelling framework adopted here is to explicitly model the observation error process, to produce a flight height distribution that accounts for the observation error. The extent to which this is successful depends on how accurately the estimates of error (95% confidence intervals on the flight height observations) reflect the true error. Sources of uncertainty and variation in the estimates of flight height are likely to be the quality of the image, and the species, size and variability in size (which in themselves may be related to age and sex), position, height and behaviour of the bird. If image quality and/or bird behaviour varies with the height of the bird, then height may also impact the magnitude of the observation error.

The digital aerial survey data record the estimated height of each bird to the nearest metre, rather than assigning birds to height categories. This higher resolution of data is likely to give substantially improved inference about detailed variation in the flight height distribution. This may be a large factor in the different distributions with the boat survey data and the digital aerial survey data. In Appendix 2 we explored error structure in the two survey types and Figures S5 and S7 show the data from different simulated surveys from the same underlying simulated distribution. In addition, by capturing images from above, digital aerial surveys may be less prone than boat based observers to missing birds which are either hidden between wave troughs or those which are flying particularly high or particularly low.

The summary of this section is that there are a number of sources of uncertainty in the estimation of flight height and not all of these are encompassed in the estimates of error. Therefore any unmodelled error or bias in the estimate of flight heights may lead to differences in the estimated distributions. The different resolution of the data types (height categories and individual height estimates) may give much richer data from the digital aerial survey data, with which to estimate the peaks and troughs of the flight height distribution.

5. Differences in the analytical methods

The analytical framework has been adapted to suit the different data types from boat and digital aerial surveys and therefore differences in the analysis may lead to differences in the flight height distribution. The Gamma mixture model performed better than the spline function for simulations tested with this Bayesian model and the digital aerial data. The spline function used in Johnston et al. (2014) may fit a different range of functions than the Gamma-mixture models used here. Alternatively, the Gamma mixture may fit better to the digital aerial survey data, given the higher resolution of the data, with a point estimate of flight height for each individual bird.

To understand the combined potential impact of 4 (differences in the accuracy of the different survey methods) and 5 (differences in the analytical methods) we used a simulation. This is described in detail in Appendix 3. We used one simulated distribution and firstly simulated sampling 500 individuals with the digital aerial data and analysed the data with the method described above. Secondly we simulated sampling almost 45000 individuals with the boat survey and analysed the data with the methods described in Johnston *et al.* (2014). This was the number of gannets assigned to height categories and used in the analysis in Johnston *et al.* (2014).

The confidence intervals from the analysis with simulated boat data are considerably larger but contain the majority of the true distribution. The credible intervals from the analysis with simulated digital aerial data are narrower, but also contain the majority of the true distribution (Appendix 3). The estimated distribution from simulated digital aerial data is more closely aligned with the true distribution. This demonstrates that a combination of different sampling and different analysis can lead to varying estimated distributions, even when the true underlying distribution is the same. However, it is not possible to separate the effects of the sampling and the effects of the analysis using this approach.

5. CONSIDERATION OF IMPLICATIONS OF DIFFERENCES IN ESTIMATED FLIGHT HEIGHT DISTRIBUTIONS

Differences in flight height distributions resulting from differing data collection methods may have implications, for example on the proportion of individuals estimated to be flying at the potential collision risk height for renewable energy structures such as wind turbines. In order to compare the proportion of individuals estimated to be flying at potential collision height as a result of the differences between flight height distribution curves from this analysis and previous work (Johnston *et al.* 2014), we used a collision risk window of 20-120m. It is recognised that this does not correspond to current turbine designs which typically have a minimum clearance height of at least 22m, but this range allows comparison with previous studies to be made. The proportion at potential collision height was calculated for each iteration of the chains (after the burn-in period and with thinning). The median of these values was calculated as the median posterior estimate of the proportion at potential collision height. The 2.5th and 97.5th quantiles were taken as the 95% posterior credible intervals for the proportion at potential collision height. These estimates and credible intervals were compared to the proportion flying at potential collision height (20-120m) estimated from distributions calculated from boat survey data (Johnston et al., 2014). Additional analyses were then undertaken to explore differences in the proportion of birds at risk of collision in relation to current turbine designs (with rotor sweeps of 22-250m, 25-253m and 30-258m above sea level). In relation to these analyses, no confidence intervals were presented for the boat-based survey data as this would have involved re-generating these distributions and was beyond the scope of this work. Consequently, for these alternative turbines, comparisons are made by considering whether the 95% CIs for the distribution derived from digital aerial survey encompass the point estimate of the proportion at risk height predicted from the distribution derived from boat-based surveys.

5.1 Comparing proportions at risk from boat and digital aerial survey data distributions

For 5/7 species, the proportion at potential collision height was not markedly different between distributions from boat and digital aerial survey data (Figure 17). For these five species either the boat survey data 95% confidence intervals included the digital aerial survey data midpoint estimate, or the digital aerial survey data 95% credible intervals included the boat survey data midpoint estimate (Figure 17; Table 2). Sandwich tern and Gannet appeared to have substantially higher estimates of the proportion at potential collision height based on the digital aerial survey data in this analysis (Figure 17; Table 2). We discuss possible reasons for this difference in these species in section 5.4. However, it should be noted that it is premature at this stage to draw firm conclusions regarding differences in the proportions at potential collision height due to differences in when and where the data were collected and differences in the sample sizes underlying the distributions.

Recent studies using tagging data have suggested that flight heights can vary depending on whether birds are actively foraging or commuting between breeding colonies and foraging areas during the summer months (Corman & Garthe, 2014 Cleasby *et al.*, 2015). For Gannet, boat-based distributions were derived from 27 different sites gathered throughout the year, whilst distributions from digital aerial survey data were derived predominantly from a single site in August with supplementary data from three other sites which gathered some data between the months April to July (Appendix 1). Consequently, the distributions derived from boat-based survey data are likely to capture a greater range of behaviour than those derived from digital aerial surveys, which were at five sites during a single year. It is therefore not possible to conclude if these differences reflect true site and seasonal differences or are an artefact of methodological differences. Since Gannets have been shown to fly at greater altitudes when foraging (Cleasby *et al.* 2015), a reason for the difference in the proportion at risk estimated from different platforms may be that data collected during digital aerial surveys are predominantly of foraging individuals although this is only conjecture at this time.

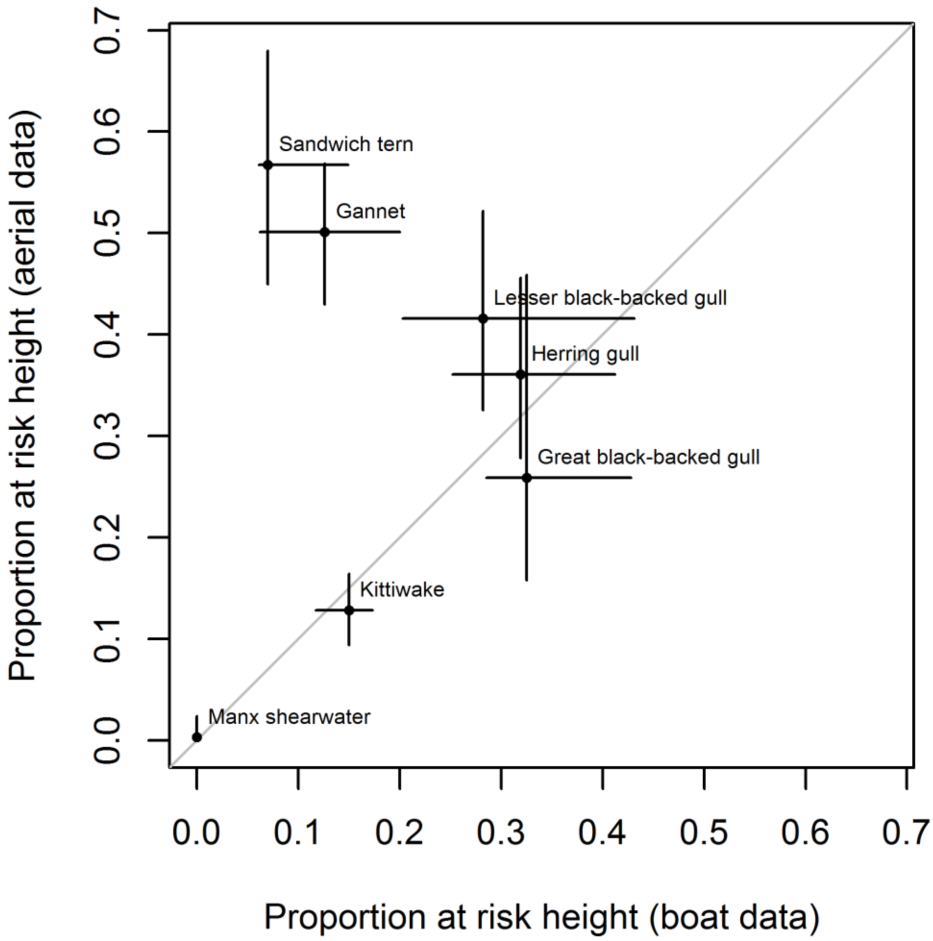

Figure 17. Estimated proportion at potential collision height (20-120m rotor swept area) from flight height distributions calculated from boat and digital aerial survey data. Lines indicate 95% confidence intervals and 95% credible intervals, respectively.

This pattern was similar when considering the three alternative heights for potential collision risk. At 22-250m (Figure 18) the patterns were similar to 20-120, with gannet and Sandwich tern showing the greatest deviation across the two distributions. As the lower limit increase, the estimated proportion at risk for Sandwich tern from the digital aerial data, reduced towards that estimated from the boat data. However, it should be noted that the majority of observations for this species are from a single site. Gannet remained the species with the greatest difference between the two estimates at all risk heights (Figures 17-18).

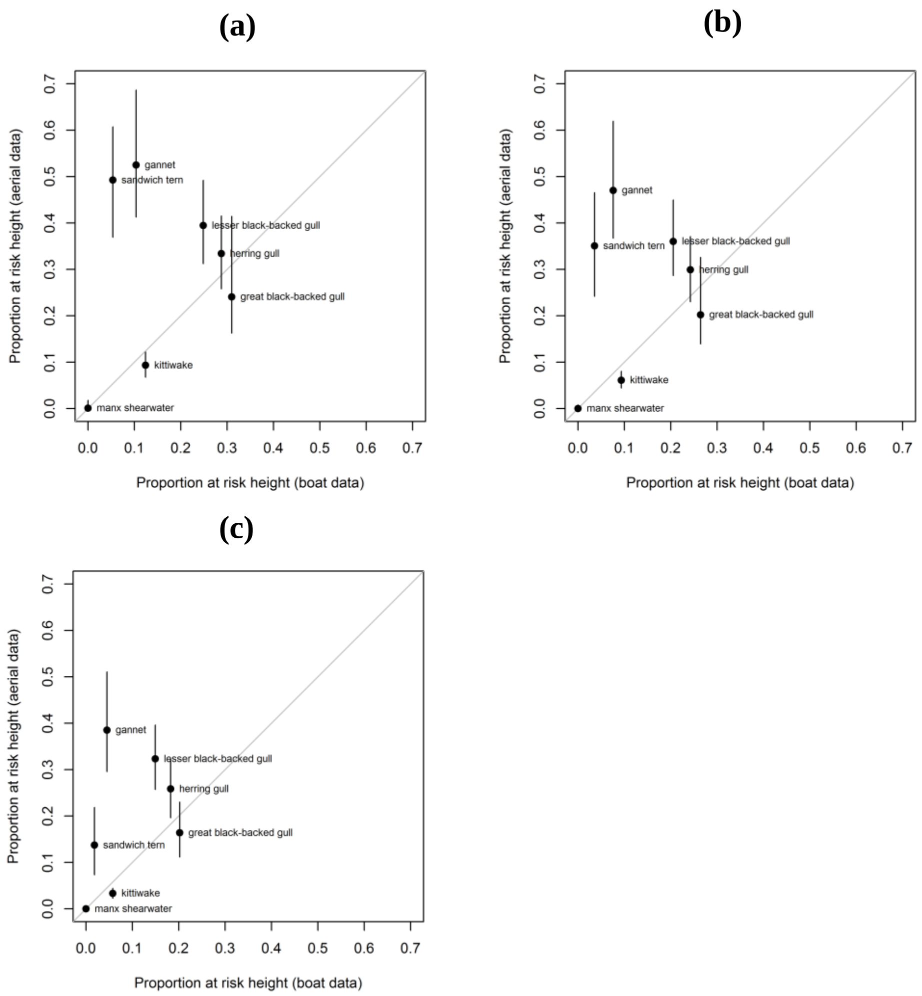

Figure 18. Estimated proportion at potential collision height from flight height distributions calculated from boat and digital aerial survey data using rotor sweeps of a) 22m-250m, b) 25m-253m, c) 30m -258m. Lines indicate 95% credible intervals and there are no confidence intervals for the estimates from the boat distributions

Table 2. Estimated proportion of birds flying at potential collision height (20-120m) and 95% credible intervals. These estimates are calculated from the above models for digital aerial survey data. Species in bold have overlap between the 95% confidence intervals and 95% credible intervals from the two models.

Species	**Proportion of birds flying at 20-120m above sea level**	
	Estimated from boat survey data (95% confidence interval)	**Estimated from digital aerial survey data (95% credible intervals)**
Kittiwake *Rissa tridactyla*	**0.15 (0.12, 0.17)**	**0.13 (0.09, 0.16)**
Lesser black-backed gull *Larus fuscus*	**0.28 (0.20, 0.43)**	**0.42 (0.33, 0.52)**
Great black-backed gull *Larus marinus*	**0.33 (0.29, 0.43)**	**0.26 (0.16, 0.46)**
Manx shearwater *Puffinus puffinus*	**0.00 (0.00, 0.00)**	**0.00 (0.00, 0.02)**
Gannet *Morus bassanus*	0.13 (0.06, 0.20)	0.50 (0.43, 0.57)
Herring gull *Larus argentatus*	**0.32 (0.25, 0.41)**	**0.36 (0.28, 0.46)**
Sandwich tern *Sterna sandvicensis*	0.07 (0.06, 0.15)	0.57 (0.45, 0.68)

5.2 Comparing kittiwake proportions at potential collision height across different sites

Flight height distributions of Kittiwake across separate sites led to substantially different estimates of the proportion of birds at potential collision height (Table 3). Although this could equally be the case for other species, kittiwake was the only species where analysis at individual sites was possible due to sample sizes. Sites 2 and 3, which are further offshore, showed some similarity in the distributions derived using digital aerial survey whilst site 1 differed markedly from both. In addition, Kittiwake numbers peaked between January and December in sites 2 and 3, but site 1 was only surveyed during August. These differences between datasets gathered using the same survey platform suggest that there are spatial and/or seasonal patterns in kittiwake flight heights. However, with the available data, it is not possible to separate out site from seasonal differences in flight height distribution. This appears to emphasise the high degree of variability in flight height data and suggest that it is difficult to distinguish between variability across the sites from variability across time. The validation exercise carried out as part of the analysis of boat survey data (Johnston *et al.* 2014) also revealed site-specific differences, since modelled distributions were often a poor match when applied to novel sites with only 44% of kittiwake data from independent sites falling within the 95% CIs of the distributions. Therefore the only firm conclusion one can reach for kittiwake on the basis of this analysis is that their flight height can be variable in space and/or time.

Table 3. Estimated proportion of Kittiwakes flying at potential collision height (20-120m) and 95% credible intervals. These estimates are calculated from the above models for digital aerial survey data, using separate analysis for three sites.

Site (sample size)	Proportion at potential collision height (95% credible interval)
Site 1 (131)	0.04 (0.01, 0.09)
Site 2 (603)	0.16 (0.09, 0.26)
Site 3 (462)	0.31 (0.19, 0.44)

5.3 Considerations for estimates of collision risk

This study found clear differences between the estimated distributions from boat and digital aerial survey data (Section 4). However, the implications for estimates of collision risk are a little more subtle. The differences in collision risk are only affected by differences in the distributions above the lower limit of the turbine blade, which here we set at 20m above sea level to compare directly to Johnston et al. (2014). So some species have distributions which look very different (e.g. kittiwake; Figure 11), but the distributions are similar above 20m and therefore give a similar estimate of the proportion flying at potential collision height (Figure 17; Table 2).

Sandwich tern and Gannet have substantially higher estimates of the proportion at potential collision height from the digital aerial survey data distributions, compared to the boat survey data distributions (Figure 17; Table 2). This difference may be attributed to several, or all, of the five factors mentioned in Section 4.4, which may be contributing to differences in the overall distributions between boat survey data and digital aerial survey data. However, due to the lack of comprehensive flight height data from every site in every month and spanning a large enough sample of years and sites from both survey platforms, it has not been possible to determine the degree to which these apparent differences are due to differences in the survey method (with perhaps one method yielding a better reflection of reality than the other) (factors 3-5 listed in section 4.4) or due to differences in the sites, periods of the year, periods of the day and weather conditions under which the data from the two platforms were gathered (factors 1 & 2 listed in section 4.4), all of which may lead to real differences in the height at which birds fly. For some species in particular, we believe that behavioural differences of birds around boats and planes (Borberg et al., 2005; Buckland et al., 2012; Spear et al., 2004; Thaxter and Burton, 2009) may be a major cause of the differences in estimates of proportion at potential collision height. Whilst data from digital aerial survey offer greater precision, in some circumstances, flight height data may only be available from boat-based surveys.

The key considerations that can be drawn from this work therefore are that:

1) Careful consideration should be given to the most suitable survey platform to collect flight height and behavioural data, with aerial survey appearing to provide a better precision overall (but with caveats relating to the factors listed in section 4.4).
2) Since flight heights of many species vary considerably both temporarily and spatially, this variability should be captured by collision risk modelling, either through use of simulation approaches like Masden (2015), or by using the upper and lower confidence intervals surrounding estimated flight height distributions in the extended Band collision risk model.

Some recommendations and next steps are outlined in Section 6.

6. CONCLUSIONS, RECOMMENDATIONS AND NEXT STEPS

6.1 Conclusions

Here we presented a methodology for deriving continuous flight height distributions from digital aerial survey data, suitable for use in the extended Band (2012) collision risk model. Digital aerial surveys estimate the height of each individual bird (with an associated estimate of uncertainty) and these data were used to estimate flight height distributions. We compared these distributions to those produced from boat data (Johnston *et al*. 2014) and found that the distributions from digital aerial data differ from the distributions from boat data. For many species, the distributions from digital aerial data had peaks above the sea surface and the mixture model selected more than one peak. In contrast, the previous distributions derived from boat data (Johnston *et al.* 2014) generally had the most parsimonious model with a peak at the sea surface and an exponential-type curve towards higher heights. We discussed several potential reasons for the differences in the estimated distributions from the two types of data and these are: 1.) differences in the flight height distributions between sites, 2.) differences in the flight height distributions between times or seasons in which data are collected, 3.) differences in bird response behaviour with the different survey platforms, 4.) differences in the accuracy of the different survey methods, 5.) differences in the analytical methods. However, due to the lack of comprehensive data from both survey platforms it was not possible to conclude with certainty why differences were observed.

For illustrative purposes the different flight height distributions from the two types of data were used to estimate the proportion at potential collision risk height. For five species the estimated proportion at risk was similar from the distributions derived from the two types of data. This suggests that despite the differences between the overall distributions, the tails of the distributions (above risk height) are sufficiently similar to lead to similar proportions at potential collision risk height for these five species. Gannet and Sandwich tern had higher estimates of the proportion at potential collision risk height from the distribution derived from digital aerial data. However our analyses suggest there is spatial and temporal variation in flight height behaviour. This result is consistent with findings for gannet and lesser black-backed gull from tagging data (Corman & Garthe, 2014; Cleasby *et al.,* 2015) and Johnston *et al*. (2014) that found evidence of site-specific variation across many species from boat data. All of this suggests that the flight heights of many species vary considerably in space and time, irrespective of the survey platform used to study them. In the face of such variation, CRM should be based on either i) generic flight height distributions derived from data pooled across many sites, months and years, and exploring the uncertainty and variation within those data, or ii) very robust site specific data covering the entire period of the year for which a species is present at a site, with multiple years of data, and again acknowledging uncertainty and variation in the data. In either case, it is likely that use of seasonal variation in the flight height figures used to derive PCH or distributions will yield a more reliable estimate of year-round collision mortality than one based on an average value of PCH or average distribution that is applied across every month.

6.2 Recommendations

For most species, the distributions derived from boat and digital aerial surveys result in similar predictions in the proportion of birds flying at collision risk height. However, the different underlying shapes of the distributions may lead to significant differences in the number of collisions predicted when using an extended version of the Band model to carry out CRM. It is important to note that if the extended Band model is used, the flight height distributions may not be transferable across platforms, i.e. distributions derived from digital aerial survey data should not be used with densities derived from boat-based surveys and *vice versa*. Digital aerial surveys detect significantly fewer birds in flight than boat surveys (WWT Consulting, 2012). Therefore, if estimates of the number of birds in flight derived from boat surveys are used with flight height distributions estimated from digital aerial survey data, may significantly over-state collision risk.

It is also clear from these results that there are notable seasonal or site-specific differences in the proportion of kittiwake at potential collision height. It is likely, that with additional information from a range of sites and across multiple seasons, the same would be found to be true for several of the other species considered here. This highlights the importance of collecting robust site-specific data where possible, to inform collision risk models. This recommendation is subject to the caveat that site-specific data must be robust in terms of numbers of birds recorded and the times of day, months of the year and weather conditions covered to reliably capture the full range of flight height behaviour exhibited by the birds over time at the site in question. Furthermore, it is recommended that provided sufficient data are available to do so, month to month variation in flight heights should be incorporated in CRM in the same way as variation in bird density is already dealt with.

This report also found seasonal differences in estimated heights for some species. Therefore we recommend that the seasonal impacts on height be considered when examining the estimated proportion at risk.

Distributions of flight heights derived from different survey platforms have different strengths and weaknesses (Thaxter *et al.* 2015) and it is important that these are considered when using distributions derived from either platform. Whilst there are some similarities in the proportion of birds at estimated risk height from different platforms, there is also evidence of seasonal and site-specific differences for some species. By making better use of site-specific data, where available in sufficient quantity to yield reliable estimates of average figures and of variance around those, it will be possible to reduce the assumptions underpinning estimated flight height distributions resulting in more robust estimates of the number of birds at risk of collision from collision risk models.

6.3 Next steps

There is a clear need to fully understand the biases and errors associated with flight height estimates from boat and digital aerial survey data. This could be done as part of a multi-platform study in which estimates of flight height are made from using both boat and digital aerial survey in the same area over a similar time frame. If possible, these data could be compared with flight height data obtained from tagged birds in the same area.

More generally, there is a need for wider collection of flight height estimates covering a range of different habitats, locations, times and species. This will help us understand the extent of variation in flight heights in different sites, years, seasons and weather conditions. A key requirement is to understand how much site specific data might be needed to capture flight height variability at a site and how flight height data should be used within the collision risk modelling. For example, whether flight heights should be stratified by month or season, and how many years of site-specific data are needed to capture inter-annual variation. This is a key requirement to inform decisions about whether site specific flight height data are adequate to provide an accurate representation of flight behaviour and hence collision risk potential at a site, or whether it would be more appropriate to use generic data on flight heights derived from a larger sample of data pooled across a range of sites, years and months.

Acknowledgements

Thanks to Natural England and The Crown Estate for providing funding for this work, and to Emily Coleman and Penny Mitchell (BTO) for managing the contract with Natural England and The Crown Estate. Thanks to Mark Rehfisch, Roger Buisson & Simon Warford (APEM) and Andy Webb, Rhys Hexter & Kit Hawkins (HiDef) for help in obtaining and interpreting the data. This work was overseen by a project steering group chaired by Tim Frayling (Natural England) and including: Orea Anderson (JNCC), Jessica Campbell (The Crown Estate), Marcus Cross (ScottishPower Renewables), Alastair Mackay (Fugro), Aly McCluskie (RSPB), Sue O'Brien (JNCC), Mark Rehfisch (APEM), Alex Robbins (SNH), Kit Hawkins (HiDef), Andy Webb (HiDef) and Jared Wilson (Marine Scotland).

References

Band, W., 2012. Using a collision risk model to assess bird collision risks for offshore windfarms (Report commissioned by The Crown Estate Strategic Ornithological Support Services (SOSS)).

Band, W., Madders, M., Whitfield, D.P., 2007. Developing field and analytical methods to assess avian collision risk at wind farms, in: Birds and Wind Farms: Risk Assessment and Mitigation. Quercus Edicions, Madrid, Spain, pp. 259–275.

Borberg, J.M., Ballance, L.T., Pitman, R.L., Ainley, D.G., 2005. A test for bias attributable to seabird avoidance of ships during surveys conducted in the tropical pacifi. Mar. Ornithol. 33, 173–179.

Bradbury, G., Trinder, M., Furness, R.W., Banks, A.N., Caldow, R.W.G., Hume, D., 2014. Mapping seabird sensitivity to offshore wind farms. PLoS ONE 9, e106366.

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.

Camphuysen, K.J., 2011. Northern gannets in the North Sea: foraging distribution and feeding techniques around the bass rock. Br. Birds 104, 60.

Camphuysen, K.J., Fox, A.D., Leopold, M.F., Petersen, I.K., 2004. Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the U.K.: a comparison of ship and aerial sampling methods for marine birds, and their applicability to offshore wind farm assessments (No. BAM - 02-2002). NIOZ Commisioned by Cowie Ltd.

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. DOI: 10.1111/1365-2664.12529

Cook, A.S.C.P., Johnston, A., Wright, L.J., Burton, N.H.K., 2012. A review of flight heights and avoidance rates of birds in relation to offshore wind farms (Strategic Ornithological Support Services SOSS-02. BTO Research Report No. 618). BTO, Thetford.

Corman, A, Garthe, S. 2014, What flight heights tell us about foraging and potential conflicts with wind farms: a case study in Lesser Black-backed Gulls (Larus fuscus). J. Ornithol. 155, 1037-1043

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

Desholm, M., Kahlert, J., 2005. Avian collision risk at an offshore wind farm. Biol. Lett. 1, 296–298.

Drewitt, A.L., Langston, R.H.W., 2006. Assessing the impacts of wind farms on birds. Ibis 148, 29–42.

Furness, R.W., Wade, H.M., Masden, E.A., 2013. Assessing vulnerability of marine bird populations to offshore wind farms. J. Environ. Manage. 119, 56–66.

Garthe, S., Hüppop, O., 2004. Scaling possible adverse effects of marine wind farms on seabirds: developing and applying a vulnerability index. J. Appl. Ecol. 41, 724–734.

Grecian, W. J., Inger, R., Attrill, M. J., Bearhop, S., Godley, B. J., Witt, M. J., Votier, S. C. 2010. Potential impacts of wave-powered marine renewable energy installations on marine birds. Ibis. 152, 683-697

Hudson, A.V., Furness, R.W., 1989. The behaviour of seabirds foraging at fishing boats around Shetland. Ibis 131, 225–237.

Irons, D.B., 1998. Foraging area fidelity of individual seabirds in relation to tidal cycles and flock feeding. Ecology 79, 647–655.

Jasra, A., Holmes, C.C., Stephens, D.A., 2005. Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. Stat. Sci. 20, 50–67.

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., Thaxter, C.B., Austin, G.E., Cook, A.S.C.P., Humphreys, E.M., Still, D.A., Mackay, A., Irvine, R., Webb, A., Burton, N.H.K., 2015. Modelling the abundance and distribution of marine birds accounting for uncertain species identification. J. Appl. Ecol. 52, 150–160.

Krijgsveld, K.L., Fijn, R.C., Japink, M., van Horssen, P.W., Heunks, C., Collier, M.P., Poot, M.J.M., Beuker, D., Dirksen, S., 2011. Effect studies offshore wind farm Egmond aan Zee. Flux, flight altitude and behaviour of flying birds (Bureau Waardenburg Report No. 10-219). Bureau Waardenburg, Culemborg.

Pearce-Higgins, J.W., Stephen, L., Langston, R.H.W., Bainbridge, I.P., Bullman, R., 2009. The distribution of breeding birds around upland wind farms. J. Appl. Ecol. 46, 1323–1331.

Rojek, N. A., Parker, M. W., Carter, H. R., McChesney, G. J., 2007. Aircraft and vessel disturbances to common murres *Uria aalge* at breeding colonies in central California, 1997-1999. Mar. Ornithol. 35, 61-69.

RSPB, 2015. Assessment of the ability of observers to accurately estimate bird flight heights from a survey vessel (Preliminary project report to MROG CRM sub-group).

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, 1851–1860.

Sotillo, A., Depestele, J., Courtens, W., Wincx, M., Stienen, E.W., 2014. Consumption of discards by Herring Gulls Larus argentatus and Lesser Black-backed Gulls Larus fuscus off the Belgian coast in the breeding season. Ardea 102, 195–206.

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. Mar. Ornithol. 32, 147–157.

Spear, L. B., Ainley, D. G., 1997. Flight speed of seabirds in relation to wind speed and direction. Ibis 139, 234-251

Stumpf, J.P., Denis, N., Hamer, K., Johnson, T.E., Verschuyl, J., 2011. Flight height distribution and collision risk of the marbled murrelet Brachyramphus marmoratus: methodology and preliminary results. Mar. Ornithol. 39, 123–128.

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

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 No. 666). BTO, Thetford.

Tyson, C., Shamoun-Baranes, J., Van Loon, E.E., Camphuysen, K.J., Hintzen, N.T., 2015. Individual specialization on fishery discards by lesser black-backed gulls (Larus fuscus). ICES J. Mar. Sci. fsv021.

WWT Consulting, 2012. Gannet population viability analysis: Demographic data, population model and outputs (SOSS-04). WWT Consulting Ltd., Slimbridge.

APPENDIX 1: SITE DESCRIPTIONS

Due to the sensitive nature of the data used in the modelling process, it is not possible to supply exact information describing each of the sites used in this analysis. However, in order to help understand how site-specific processes may be influencing the distributions derived by this analysis we present summary information describing each site in the following table.

Site	Minimum distance From Shore	Months Surveyed (Years Surveyed)	Species Present	Number of Each Species Recorded	Month of peak abundance
1	0 km	August (1 Year)	Gannet	82	August
			Great Black-backed Gull	3	August
			Herring Gull	15	August
			Kittiwake	131	August
			Manx shearwater	9	August
			Sandwich Tern	19	August
2	WITHELD	January – December (2 Years)	Great Black-backed Gull	104	December
			Herring Gull	59	January
			Kittiwake	603	December
			Lesser Black-Backed Gull	42	November
3	WITHELD	January – December (2 Years)	Great Black-backed Gull	146	January
			Herring Gull	74	January
			Kittiwake	462	January
			Lesser Black-backed Gull	66	January
4	15 km	April – July (1 Year)	Gannet	25	July
			Great Black-backed Gull	3	June
			Herring Gull	3	July
			Kittiwake	100	July
			Lesser Black-backed Gull	13	July
			Manx Shearwater	223	June

Site	Minimum distance From Shore	Months Surveyed (Years Surveyed)	Species Present	Number of Each Species Recorded	Month of peak abundance
5	0 km	April – July (1 Year)	Gannet	12	June
			Great Black-backed Gull	16	June
			Herring Gull	30	June
			Kittiwake	16	April
			Lesser black-backed Gull	36	June
			Sandwich Tern	181	June
6	8 km	April – June (1 Year)	Gannet	35	June
			Herring Gull	19	June
			Kittiwake	62	April
			Lesser black-backed Gull	5	June
			Manx shearwater	4	April
7	27 km	March – December (1 Year)	Lesser Black-backed Gull	141	September
8	0 km	August (1 Year)	Gannet	69	August
			Great Black-backed Gull	8	August
			Herring Gull	22	August
			Kittiwake	55	August
			Lesser black-backed Gull	3	August
			Manx shearwater	4	August
			Sandwich Tern	5	August

APPENDIX 2: SIMULATIONS

To understand how well the modelling process can recreate the underlying distribution, we simulated data and analysed these data for which the underlying true distribution is known. Simulated data were generated using three different distributions – a log-Normal mixture, a Gamma mixture, and a spline. All of these generating distributions are non-parametric distributions and therefore are reasonably realistic distributions of seabird flight heights and a range of distributional shapes are enabled. The three generating distributions are described in detail below and there are also examples of random distributions provided.

Spline Generating Distribution

To simulate a spline distribution, 4 distributions were randomly generated with a uniform distribution, a log-normal distribution and two Gamma distributions. A number of individuals (birds) were selected from each of the distributions and these numbers were generated randomly from a Poisson distribution with mean 1000. Some of the parameters of each of the distributions were also randomly generated. The four distributions were defined as:

$$X_1 \sim U(0, 150)$$

$$X_2 \sim \exp(\ N(\ \mu = N(2, 0.5), \sigma = 1.5))$$

$$X_3 \sim Gamma(shape = Pois(1, 10), rate = Gamma(2, 1))$$

$$X_4 \sim Gamma(shape = Pois(1, 5), rate = Gamma(1, 1))$$

The draws from these four random distributions were combined and a spline fitted through the log of flight heights. This spline was converted to a probability distribution by dividing by the sum of all fitted values between 0-300m. This spline-based probability distribution was used to produce non-parametric simulated distributions (Figure S1).

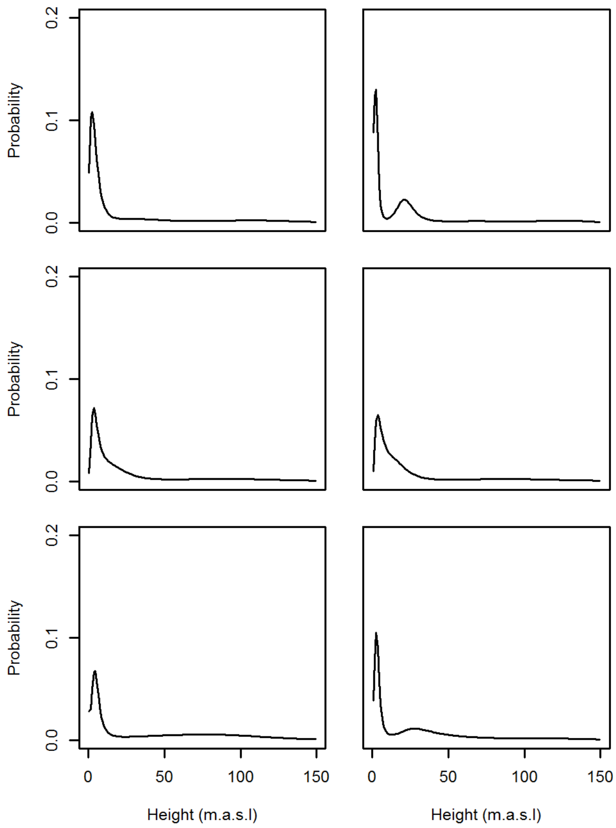

Figure S1. Example randomly generated probability distributions using a spline function.

Simulating the Observations

For each simulation, 500 birds were randomly simulated from a distribution generated from one of the three generating distributions. Bird flight heights are not measured directly, but derived and calculated from photo characteristics. Flight heights are therefore observed with error and it was important to simulate an observation process. Initially to assess how well each analysis distribution is able to fit to the simulated distributions, we used a small observation error. The observation process for each bird was assumed to be an independent error distribution and flight heights were assumed to be observed with no bias (i.e. centred on the true value). The simulated error distribution was assumed to be normally distributed and was defined as:

$$observed\ height_i \sim N(true\ height_i,\ \sigma_i^2)$$

So the observed height of an individual bird *i* was estimated from a normal distribution centred on the true height and with an individual standard deviation for bird *i*. The standard deviations were drawn from a Gamma distribution with a shape parameter 7 and a variable rate parameter. To achieve an average standard deviation of 5m, we used a rate parameter of 1.4.

$$\sigma_i^2 \sim Gamma(7, rate)$$

This produced a distribution of standard deviations with a desired median standard deviation (Figure S2). Comparisons of simulations with large and small error distributions will help clarify the contribution of error in the observation process to the estimated distributions.

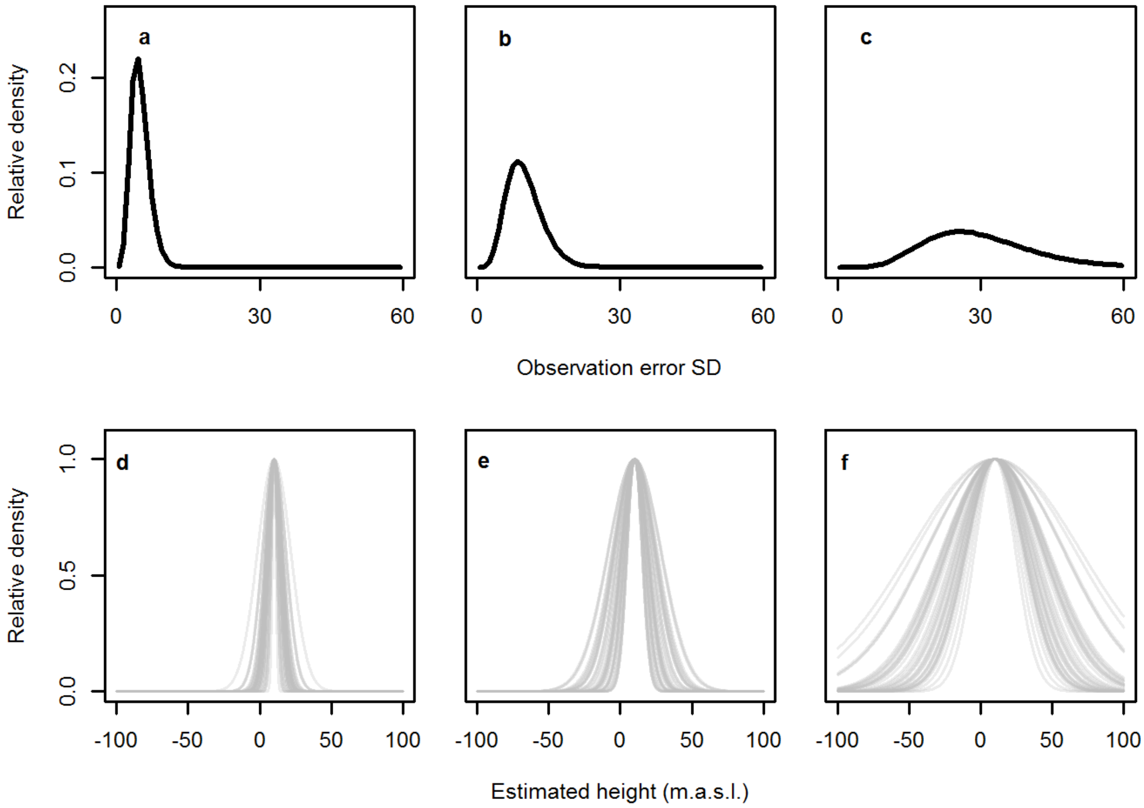

Figure S2. Example distributions of the observation standard deviation with a median SD of a) 5m; b) 10m and c) 30m. Examples of 40 random error distributions for a hypothetical bird flying at a height of 10m above sea level, with median standard deviations of d) 5m; e) 10m and f) 30m.

For the first 100 random simulations we used a standard deviation resulting in an average 95% confidence range in the error distribution of 5m. This simulates a flight height estimation process which has good precision. The second 100 random simulations we used a larger standard deviation, such that the average 95% confidence range was 30m. This simulates a flight height estimation process with less precision. Examples of the effect of this observation error on the estimates of flight heights is shown in Figure S3.

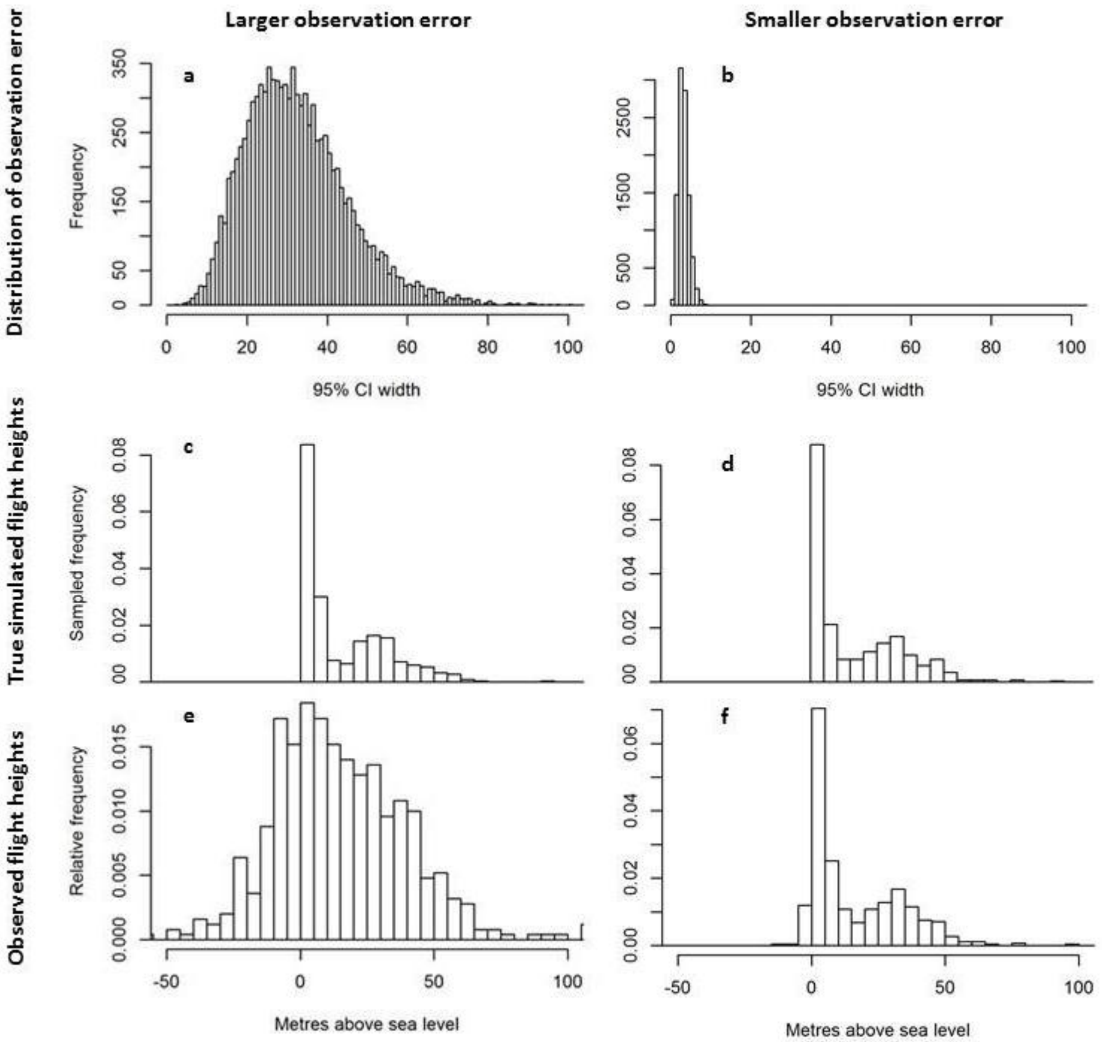

Figure S3. Two simulations were conducted: one with large observation error (left hand column) and one with small observation error (right hand column). The respective simulated flight heights (c and d) are from the same distribution, but the observed flight heights (e and f) vary substantially due to the different error distributions.

APPENDIX 3: COMPARING DISTRIBUTIONS FROM BOAT SURVEY DATA AND DIGITAL AERIAL DATA

In order to investigate the potential differences between the distributions derived from data from boat surveys and aerial surveys, we conducted a simulation. Using one simulated distribution from section 3.1 (Figure S4) we simulated surveying by digital aerial and by boat methods.

Firstly, we simulated digital aerial survey methods described in section 3.1 and Appendix 2. This simulated estimating the flight height of 500 birds with error, with the mean standard deviation of the observations at 30m (this is similar to the digital aerial survey data) (Figure S5). This was analysed as described in section 3.2 to estimate a flight height distribution (Figure S6).

Secondly, we simulated a survey protocol similar to the boat surveys. We used the boat data for gannets from Johnston *et. al.* (2014) to act as a framework for simulating the boat distributions. We simulated data collection that followed the pattern of the gannet data. We used the height bands and number of individual gannets observed at each site, to replicate the sampling regime of the boat surveys (Figures S7 & S8). We used the simulated flight height distribution and simulated sampling this using the gannet distribution of sites and number of individuals. We then analysed this simulated boat survey data with the analysis protocol for boat data (Johnston *et al*. 2014) (Figure S9).

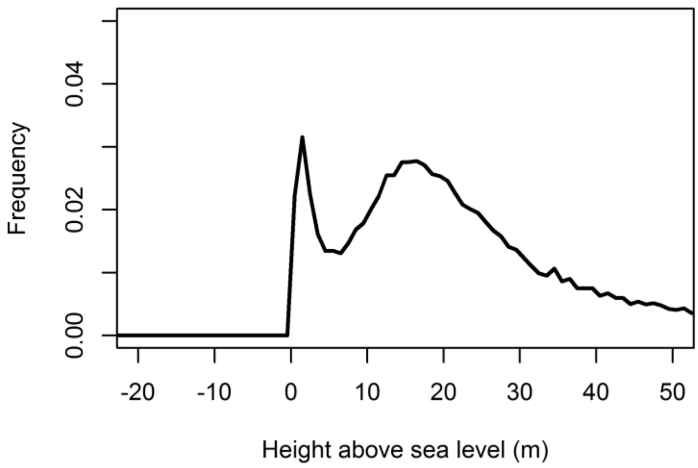

Figure S4. True distribution of flight heights in the simulation.

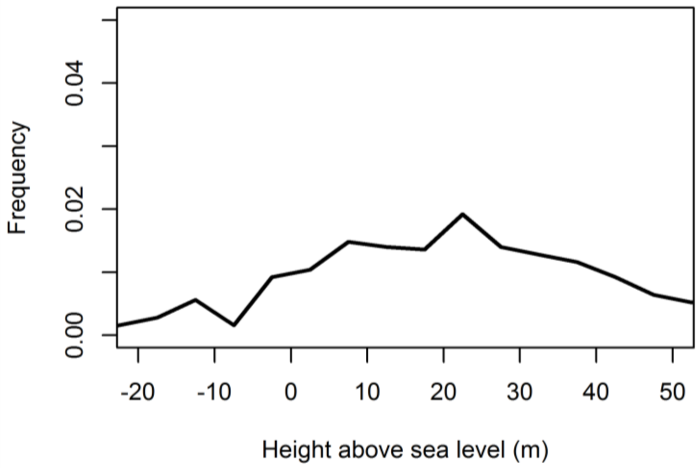

Figure S5. Distribution of estimated flight heights of 500 birds using the simulated survey protocol similar to digital aerial surveys, so heights are estimated with error.

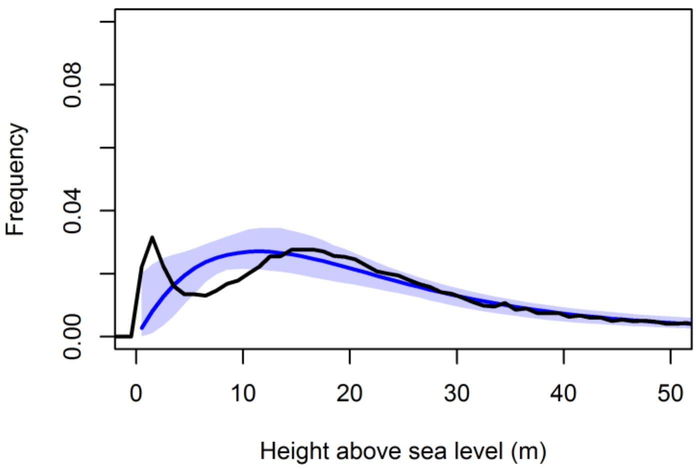

Figure S6. Estimated distribution of flight heights using the simulated survey protocol similar to digital aerial surveys and the modelling approach described above. Estimated distribution (blue line) and 95% credible interval (pale blue) compared to true simulated distribution (black line).

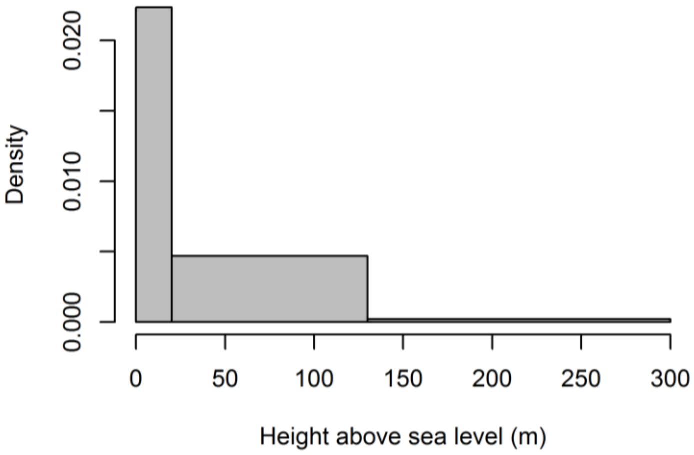

Figure S7. Distribution of estimated flight heights from one site using the simulated survey protocol similar to boat surveys, in which birds are assigned to height categories.

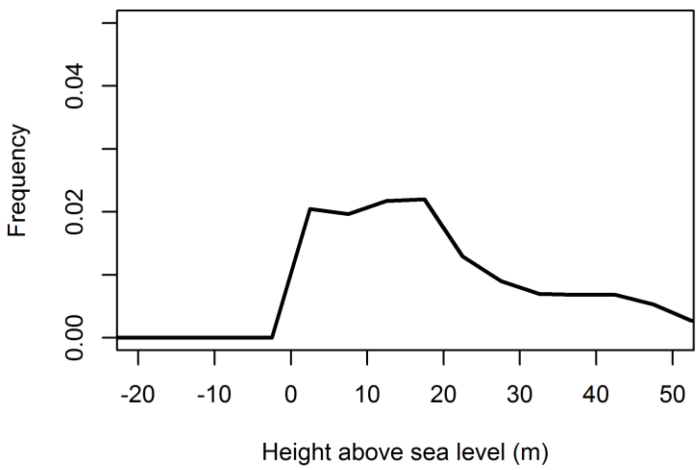

Figure S8. Distribution of estimated flight heights using the simulated survey protocol similar to boat surveys, so birds are assigned to height categories. To create this histogram, all birds within each height category are assumed to be evenly distributed across the height category (replicating the level of information available in the height categories).

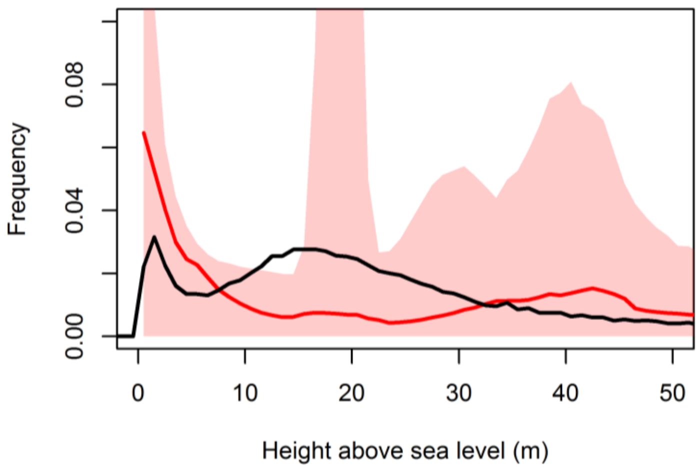

Figure S9. Estimated distribution of flight heights using the simulated survey protocol similar to boat surveys, using the modelling approach described in Johnston et al. (2014). Estimated distribution (red line) and 95% credible interval (pale red) compared to true simulated distribution (black line).

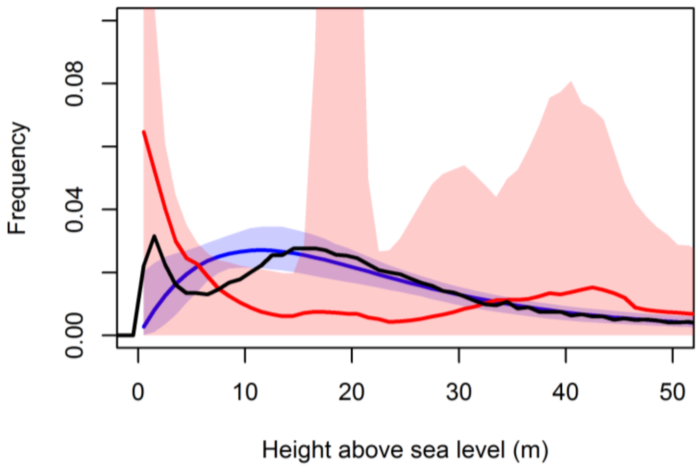

Figure S10. Comparison of true simulated distribution (black line), the estimated distribution using the approach to analyse digital aerial data described above (blue) and the estimated distribution using the approach to analyse boat survey data described in Johnston et al. (2014) (red).

These results indicate that a combination of the data collection regime and the analytical approach may lead to different estimates of the same underlying distribution. Boat surveys and digital aerial surveys have very different characteristics of error and precision. It is also important to note that birds may have different behavioural responses to each survey platform. Whilst many species show attraction to, or displacement by, boats (Camphuysen *et al.* 2004, Schwemmer *et al.* 2011) the altitude of digital aerial surveys may be less likely to trigger a response (Rojek *et al.* 2007, Buckland *et al.* 2012). These differences may influence the final derived distributions, but cannot be incorporated into simulation excercises such as this. The analysis protocols for the two sources of data use different analytical structures to estimate the flight height distributions. The combination of the data collection and the analysis may result in differences (Figure S10), however, it is impossible to separate here whether that is a consequence of the data collection or the analysis. In reality, it is likely to be a combination of both of these factors.

APPENDIX 4: UPDATED FLIGHT HEIGHT RISK SUMMARY

Here we have updated Table 34 from the Report from Objective 1 (Thaxter et al., 2015) to include estimates of the proportion at potential collision risk from the digital aerial survey flight height distributions presented here. We have calculated the median and 95% credible intervals for the proportion of the population flying between **30m and 150m** above sea level. Although different to the risk height bands used elsewhere in this paper, we have copied this directly from Thaxter et al. (2015) and therefore use the lower and upper heights used in this report. The results in this table differ from those listed in Table 2 (20-120m). Note that the information in this table is intended to illustrate how comparable estimates of flight heights from different platforms may be. However, given the data available, it is not possible to determine whether any differences are the result of the platform from which the measurements were made or if they may relate to site or seasonal differences in data collection.

Table S3. Summary of flight height distribution data from different methods for UK species, taken from (Thaxter et al., 2015). Updated with an additional column from the flight height distributions in this report. Flight height data was summarised in two ways: (1) As a percentage of the flight height distribution (for example, % time/birds/GPS fixes), at or below minimum turbine height where risk of collision is reduced – based on the studies reviewed, we assumed a vertical turbine rotor sweep zone (RSZ) of **30-150 m**; and (2) Percentage of the distribution at collision risk height using study-specific RSZs. Highlighted cells indicate a subjective gradation of risk (green = low, yellow = medium, red = high) based on these two data summaries (see key).

Species	Visual methods (+)		Tags GPS and altimeter (++)	Radar[b] (++)	Laser rangefinder[h] (+)	Digital aerial survey data
	Boat[a]	Visual Panorama[b]				
Common Eider	34.6% [3.5-55.8 CI] in RSZ			Marine duck species: <100 m, most <10m		
Common Scoter	1.9% [0.1-10.9 CI] in RSZ	*ca* 30% at RSZ outside WF				
Red-throated diver	6.2% [1.5-32.3 CI] in RSZ			Diver species: Variable, generally low <30 m ASL.		
Black-throated diver	8.1% [6.8-33.1 CI] in RSZ					
Grebe spp				Variable, generally low <50 m ASL.		
Northern Fulmar	1.0% [0.0-9.2 CI] in RSZ					
Manx Shearwater	0% [0.0-0.0 CI] in RSZ					0.0% [0-0.2 CI]
Northern Gannet	12.6% [6.2-20.0 CI] in RSZ	41% at RSZ outside WF, 21% inside	Plunge-dives, most at 11–60 m (mean±SE = 37.1±2.8 m; range 3-105 m)[c]	Most <10 m, some foraging up to 50 m searching for food. Gannets plunge from 10-30 m.	Boxplot whisker range: 1.7-40.5 m, median 18.8 m, whisker RSZ overlap 26.3%, IQR overlap 0%	34.5% [28.6 – 40.8 CI]
Great Cormorant	1.7% [0.8-27.1 CI] in RSZ	24% at RSZ outside WF, 33% inside		Cormorant species: Low-intermediate altitude most <5m not higher than 75m		
European Shag	12.6% [2.0-64.3 CI] in RSZ					
Arctic Skua	2.6% [1.7-10.0 CI] in RSZ					
Great Skua	5.9% [3.5-17.9 CI] in RSZ		<5 m; 4.4% daytime collision risk height[d]			
Black-legged Kittiwake	15.0% [11.7-17.3 CI] in RSZ	ca 50% at RSZ outside WF, ca 40% inside		Gull species: see below	Boxplot whisker range: 1-34.8 m, max ca. 80 m; median 16.6 m, whisker RSZ overlap 13.9%, IQR overlap 0%	3.3% [2.3 – 4.4 CI]
Black-headed Gull	13.9% [5.7-25.5 CI] in RSZ	41% at RSZ outside WF, 21% inside		Migrating flocks mostly in the RSZ and above; gull species see below		

Little Gull	0.0% [0.0-100.0 CI] in RSZ			Gull species: Locally < 50 m above sea level, mean (foraging, travelling during breeding) up to 250 m.	Boxplot whisker range: 8.6-24.5 m, max ca. 48 m; median 18.8 m, whisker RSZ overlap 0%, IQR overlap 0%	
Common Gull	21.9% [19.0-30.1 CI] in RSZ	46% at RSZ outside WF, 55% inside				
Lesser Black-backed Gull	28.2% [20.3-43.1 CI] in RSZ	>50% at RSZ outside/inside WF	89 % fixes below 20 m[e] Typically <20 m, most <5 m; 31.2% daytime at collision risk height [d] Estimation from figure 6: Spring and Autumn migration travel ca. >70% values <250 m AGL (coarse banding)[f] ca.90% flying fixes < 25m (Fig 5.12); 3.7% >75 m; Lesser Black-backed Gulls were more common than Herring Gulls >75 m (5.2 and 2.4%)[g]		Boxplot whisker range: 0-69.6 m; median 26.3 m, whisker RSZ overlap 56.0%, IQR overlap 37.4%; max, ca 131 m	32.3% [25.7 – 39.7 CI]
Herring Gull	31.9% [25.2-41.2 CI] in RSZ	>50% at RSZ outside/inside WF	ca.90% flying fixes < 25m (Fig 5.12); 3.7% >75 m; Lesser Black-backed Gulls were more common than Herring Gulls >75 m (5.2 and 2.4%)[g]		Boxplot whisker range: 0-74.2 m; median 32.4 m, whisker RSZ overlap 58.0%, IQR overlap 42.2%; max, ca 180 m	25.8% [19.6 – 32.3 CI]
Great Black-backed Gull	32.5% [28.5-42.8 CI] in RSZ	>50% at RSZ outside/inside WF			Boxplot whisker range: 6.1-66.1 m; median 34.4 m, whisker RSZ overlap 59.5%, IQR overlap 66.9%	16.2% [10.8 -22.6 CI]
Sandwich tern	7.0% [6.1-14.9 CI] in RSZ	ca 50% at RSZ outside WF, ca 30% inside		Tern species: Generally up to 20 m average, but through RSZ on migration		13.7% [7.3 – 21.9 CI]
Common Tern	7.4% [4.4-9.9 CI] in RSZ					
Arctic Tern	4.0% [0.6-14.3 CI] in RSZ					
Common guillemot	0.4% [0.0-10.2 CI] in RSZ			Alcid species: Hardly ever higher than 50m and nearly always very low <5m		
Razorbill	2.7% [0.0-13.7 CI] in RSZ					

Little Auk	3.6% [0.0-5.0 CI] in RSZ					
Atlantic Puffin	0.0% [0.0-6.8 CI] in RSZ					

a Johnston *et al*. 2014a; b Krijgsveld *et al*. 2011; c Garthe *et al*. 2014; d Ross-Smith *et al*. (unpubl data); e Corman & Garthe (2014); f Klaassen *et al*. (2011); g Ens *et al*. (2008); h Mendel *et al*. (2014, extracted from Fig 11.12).

Key
Green = <10% time/birds/fixes > 30 m; <10% at collision risk height (in RSZ)
Yellow = 10-30% time/birds/fixes > 30 m or <30% at collision risk height (in RSZ)
Red = >30% time/birds/fixes >30m; or more than 30% at collision risk height (in RSZ)
Grey = Hard to categorise or lacking full distribution