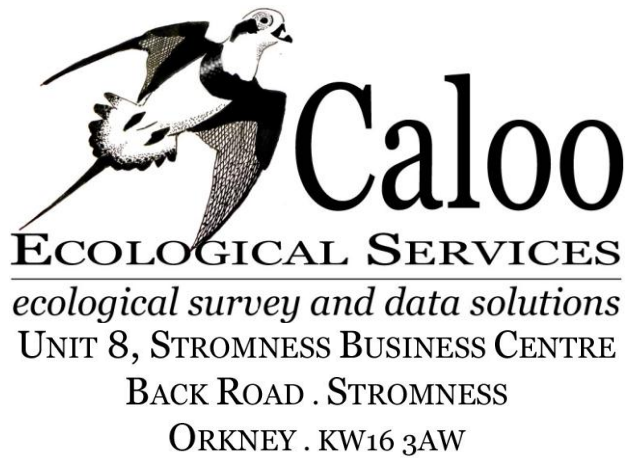


**Distance sampling analyses of ESAS survey results  
for the Brims Tidal Array Project**

**Nigel Harding  
April 2015**



---

TELEPHONE: 01856 852 814  
MOBILE: [REDACTED]  
EMAIL: [caloo.ecological.services@phonecoop.coop](mailto:caloo.ecological.services@phonecoop.coop)

## 1 Introduction

The Brims Tidal Array is a proposed tidal energy development by SSE Renewables and Open Hydro. The area being investigated for the proposed development lies to the south of the island of Hoy, off the Brims Ness headland (Figure 1). Based on present knowledge it is anticipated that a tidal array of up to 200 megawatts (MW) capacity could be installed on the site, generated by up to 200 tidal energy devices. This report describes the distance sampling analyses used to calculate the abundance and density estimates for seabirds that will underpin the environmental impact assessment for the project with respect to these species.

Distance sampling is a widely-used group of closely related methods for estimating the density and/or abundance of biological populations from data collected usually using line transects or point counts (Buckland *et al.* 2001, 2004). It caters for the fact that animals more distant from the observer are less likely to be detected, and corrects the resulting population estimates accordingly. For distance sampling to be applied, perpendicular distances of clusters of animals from the survey line must be recorded. A detection function is then fitted to these observed distances, and used to estimate the proportion of objects missed within transect. This then allows an absolute estimate of the number and density of animals present to be made. Key assumptions of the standard distance sampling methods applied here are:

- All animals on the transect line (i.e. at distance zero) should be detected.
- There should be no responsive movement prior to detection.
- Distance to animals should be measured without error.
- The detection function should have a wide shoulder (i.e. most animals should be detected out to a reasonable distance).

## 2 The Data

The data analysed here was collected using standard ESAS methodologies (Camphuysen *et al.* 2004) on 20 survey dates between June 2013 and May 2014 (Table 1a).

The proposed development lies to the South of South Walls and Cantick Head (Figure 2). At the start of the project a provisional boundary was defined for the area potentially suitable for development (Figure 2). The survey area was initially defined as all areas within a 4 km buffer around this provisional boundary (Figure 2). Areas within this 4 km buffer but to the North of South Walls and Cantick Head, were deemed very unlikely to be influenced ecologically by the proposed development, and were excluded. These areas were delimited by a straight line running between Cantick Head and the Ool on Switha. Areas excluded on this basis include North Bay, Longhope, Kirkhope and Cantick Sound. In addition, to ensure safe operation of the vessel, all areas within 500m of the shore were also excluded. Subsequently the area within which development could potentially take place was defined more precisely. This revised area potentially suitable for development almost wholly falls within the initial provisional boundary (Figure 2), with only very small areas (a total of 2.3 ha out of 1105.7 ha, 0.2%) at the site's north western and south eastern corners falling outside the initial provisional boundary. Therefore, excluding areas to the north of Cantick Head or within 500m of the shore, the survey area includes nearly all (i.e. 10601.5 ha out of 10614.22 ha, 99.88%). sea areas within 4 km of any potential development. This survey area was systematically covered by 11 transects spaced at 1.4 km intervals within a randomly positioned grid (Figure 2). With transects 0.3 km wide, and spaced at 1.4 km intervals 21.4% of the survey area would be expected to fall within the transect during each complete survey.

Table 1a presents estimates of coverage for each date. The coverage of each transect on a particular date is estimated by expressing the distance covered by the vessel (derived from global positioning system estimates of the ship's position at 1 minute intervals) as a percentage of the length of the transect. The overall coverage is estimated by expressing the total distance covered by the vessel on transect as a percentage of total transect length. As in practice the vessel can only follow the transect approximately, particularly given the vigorous tidal regime operating in the area, these percentages can sometimes be greater than 100%.

The intention was to carry out a single survey during each calendar month, with all transects being covered within a single day. On the 16<sup>th</sup> May 2013 poor sea state conditions led to the survey being abandoned with only one transect partially covered (Table 1a); data from this date has been excluded from all analyses. In February 2014, survey coverage was extended over two dates, the 17<sup>th</sup> and 19<sup>th</sup>, and these have been combined into a single survey. Table 1b presents the survey effort achieved on each survey, excluding the data from the abortive survey in May 2013, and combining the data from the two survey dates in February 2014 into a single survey, assigned to the 18<sup>th</sup> February 2014. Between late March 2012 and early March 2014 18 surveys were conducted, with complete coverage being achieved in 12 surveys, and partial coverage in 6 surveys.

For the 12 complete surveys, overall coverage across all transects always exceeded 96%, whilst coverage of individual transects was never less than 83%, and for 11 out of the 12 surveys was always greater than 89% (Table 1b).

For the 6 incomplete surveys overall coverage varied between 53% and 89% (Table 1b). On the 20<sup>th</sup> August 2012, all transects apart from transect 22 were fully covered and an overall coverage of 89% was achieved. Three of the incomplete surveys (11<sup>th</sup> December 2012, 4<sup>th</sup> March 2013, 12<sup>th</sup> March 2014) showed a similar pattern, with complete coverage of transects 2-14, but no coverage of transects 16 to 22, providing overall coverage of 53-56%. During a fifth survey on the 23<sup>rd</sup> October 2013, again transects 2-14 were completely covered, but there was only partial coverage of transects 16 and 18, and no coverage of transects 20 and 22, providing overall coverage of 68%. Thus five of the six incomplete surveys show a similar pattern, with complete coverage of the 7 easternmost transects (transects 2-14) and either partial or no coverage of the 4 westernmost transects (transects 16-22). This reflects the poorer sea conditions often encountered on these westernmost transects, reflecting their greater exposure to Atlantic swells. The sixth survey, conducted on the 17<sup>th</sup> and 19<sup>th</sup> February 2014, shows a different pattern of a coverage, with complete coverage of transects 4 - 10 and transects 20 -22, partial coverage of transects 2 and 18, and no coverage of transects 12 - 16.

Surveys took place over 25 calendar month period, from late March 2012 to early March 2014. Complete surveys, in which all transects were covered, were achieved in 10 of these 25 calendar months (March, April, May, June, and July in 2012; February, March, June, July, September in 2013). Surveys in which some but not all transects were covered were achieved in 5 months (August and December 2012, October 2013, and February and March 2014). No coverage was achieved in 10 out of the 25 months (September, October and November 2012; January, April, May, August, November and December 2013, January 2014). Although there were no surveys in April or May 2013, additional complete surveys were achieved in the adjoining months (on the 30<sup>th</sup> March 2013 and the 3<sup>rd</sup> June 2013) so as to maintain as even coverage throughout the year as closely as possible. In March 2013, in addition to complete surveys conducted on the 5<sup>th</sup> March and 30<sup>th</sup> March, the latter in lieu of the April survey, an incomplete survey was completed on the 4<sup>th</sup> March. When presenting the results for individual surveys the results from this partial survey have been included, to illuminate variation between surveys within a month. However, when calculating seasonal

abundance estimates, the results from this additional incomplete survey on the 4<sup>th</sup> March have been excluded, to make the sample more representative.

Difficult sea conditions associated with the strong tidal currents in the survey area prevented surveys in some months. In other months, although surveys were conducted, poor sea conditions meant it was often not possible to complete all transects. In particular, completing the westernmost transects was often difficult. The incomplete coverage achieved in these partial surveys means that the density of birds encountered during them may not be representative of the densities across the whole survey area. Thus, including the results from these incomplete surveys will potentially lead to bias in our density estimates. However, excluding them would carry a heavy cost in terms of achieving representative cover throughout the year. Thus, results from these incomplete surveys have been retained in the analyses. However, the potential for bias in their results as a result of incomplete coverage should be borne in mind when interpreting the results. As noted previously a common pattern across the incomplete surveys was to achieve complete coverage of the 7 easternmost transects (transects 2-14), but partial or no coverage of the 4 easternmost transects (transects 16-22). One of the analyses we perform below, for the 12 surveys where complete coverage was achieved, is to compare density estimates for each species based upon the 7 easternmost transects with the corresponding estimates based upon the four easternmost transects during the same survey. This should allow us to better understand any potential biases introduced into density estimates because of incomplete coverage.

On the first two surveys a small amount of survey effort (2.1 km (2.7%) on the 27/03/2012 and 3.6 km (4.6%) on the 18/04/2012) was conducted in sea state 5, whereas Camphuysen *et al.* (2004) recommend that all surveys should be conducted in sea states of four or less. Data from this survey effort outside standard conditions has been retained in the analyses to retain full coverage of the site on these two dates. The inclusion of sea state as a covariate in the detection function modelling (see below) should further minimise any bias including this small amount of survey effort conducted outside standardised survey conditions has on density and abundance estimates.

Initially we derive density estimates for each species on each survey. From these we derive seasonal abundance estimates for each species. The seasons used for each species are specific to that species and reflect its local phenology. Our assessment of impacts is based upon seasonal abundance estimates rather than monthly abundance estimates, as we think this provides a more reliable basis for assessment. We derive seasonal abundance estimates both for separate years and also across years. Generally, each season includes several months. Therefore even, for seasonal estimates within a year, each seasonal abundance estimate is based upon several surveys. Given the variation between survey dates within a season in abundance and density the much larger sample size (in terms of survey dates) upon which seasonal abundance estimates are based should make them a much more reliable basis for impact assessment than would using monthly estimates. Furthermore, as the density estimate for each season within a year is based upon multiple surveys, and so has a measure of error attached, we can potentially identify year to year variation in the density of birds recorded within a particular season. Given the large temporal variation to be expected in the density of a seabird species within a month a single survey within a month cannot be assumed to be representative. Furthermore, with only a single survey, there is no measure of the variation between different survey dates within a month. Thus, although confidence limits can be attached to the density estimates from a particular survey on a particular day within a month, we cannot compute reliable confidence limits for the expected density of birds across the month. Even across two survey years, a single survey in each month in each year provides a sample size of just two surveys for each month. With a sample size of just two surveys the mean is still unlikely to provide a reliable best estimate of the expected density of birds in that month, and any measure of the variance is also likely to remain unreliable. Furthermore, with only one survey in a month in

a particular year, year to year variation in density cannot be separated from day to day variation. Thus, for these reasons we conclude seasonal abundance estimates should provide a sounder basis for impact assessment than would monthly abundance estimates. A further advantage of using seasonal abundance estimates is it potentially provides a basis for estimating density even across those months for which no survey effort was possible, whether within a single year (September, October and November 2012; January, April, May, August, November and December 2013, January 2014) or across years (January, November, December across both survey years).

For birds in flight, the survey data upon which abundance and density estimates are based consists of birds recorded on effort and in transect during the snapshot counts (Camphuysen *et al.* 2004). For birds on water, the survey data, which is used to estimate the detection probabilities as well as abundance and density, consists of all sightings of birds on the water on effort and within transect. In defining which observations were in transect, observations outside the survey area, or beyond the ends of the transect, or when the vessel was not travelling along the transect have been excluded. These strict criteria have been adopted to ensure that across the whole project exactly the same underlying data are used as the basis for different types of analysis. Tables 2 to 5 present the following summary statistics for each species, separating birds in flight from birds on the water:

- Table 2 presents the number of sightings for each survey, and in total across all surveys.
- Table 3 presents the total number of individual animals recorded, for each survey, and in total across all surveys.
- Table 4 presents the mean number of animals in each sighting (i.e. mean cluster size) for each survey.
- Table 5 presents the median number of animals in each sighting (i.e. median cluster size) for each survey.

### **3 Software**

All analyses were carried out using programmes written in R (R version 3.0.1 (2013-05-16), R Core Team 2013) with the distance sampling analyses performed using functions from the *mrds* library (Laake *et al.* 2012).

Although the majority of the manipulation of spatial data has been carried out using the libraries available within R (Bivand *et al.* 2008), ESRI Arcview 9.3 was used for some tasks (e.g. buffering of site boundaries, presentation of some maps).

### **4 Detection Function Modelling**

The purpose of detection function modelling is to estimate the proportion of animals observers fail to detect, so that estimates of density and abundance can be corrected accordingly. For ESAS data, detection function modelling is only possible for birds on the water as no distance data are recorded for birds in flight. For birds in flight, we have assumed a probability of detection of 100% in the 300m x 300m snapshot recording box.

For birds on the water, the ESAS methodology records birds into five distance bands A-E (0-50m, 50-100m, 100-200, 200-300m, 300m+). As no distance data are available for sightings beyond 300m (distance band E), this data cannot be included in the analyses, so that our detection function modelling could only be based on four distance bands at most. Four distance bands is the absolute minimum for detection function modelling (Buckland *et al.* (2001:262), so that further truncation to remove outliers, or further grouping of data into a smaller number of distance intervals to overcome potential problems such as heaping, errors in distance measurement or evasive movement prior to detection (Buckland *et al.* 2001), were not available as analysis options.

To estimate detection probabilities we use as our dataset all observations of birds on the water within transect, across all surveys from late March 2012 to early March 2014 inclusive (Table 1b). We model the probability of detection separately for common and rare species. Common species are defined as those with 30 or more observations of birds on the water across all surveys, and rare species as those with less than 30 observations (Table 2a). This definition yields eight common species (fulmar, gannet, shag, great skua, kittiwake, common guillemot, razorbill and puffin) and thirteen rare species (red-throated diver, black-throated diver, great northern diver, Manx shearwater, storm petrel, cormorant, Arctic skua, common gull, herring gull, great black-backed gull, Arctic tern, black guillemot and little auk). The definition of common and rare species used is consistent with the advice of Maclean *et al.* (2009), in the context of fitting separate detection function models for each species, that distance sampling analysis should only be applied to species with 30 or more records and that otherwise the probability of detection should be estimated on the basis of the generic JNCC correction factors published in Stone *et al.* (1995). Thirty observations is considerably less than the 60-80 observations recommended by Buckland *et al.* (2001:228) as the minimum required for reliable fitting of the detection function.

#### 4.1 Detection function modelling for common species

For common species, we fit a single detection function across all species. Variation in the probability of detection between species is captured by including species as a covariate in the model, with sightings for all species with less than 30 observations combined into a single “Other species” category. The shape of the detection function is modelled as a half normal key function with no adjustment terms (Buckland *et al.* 2001).

Detection functions are fitted on the assumption that it is the sightings as recorded in the field (e.g. 3 guillemots in a group) that are independently detected rather than the individuals within these clusters. This should make the fitted detection functions more reliable as a relatively small number of clusters holding high numbers of individuals will not potentially have undue influence. Furthermore, it should also avoid obtaining spuriously high estimates of precision by avoiding overestimating the number of independent observations underpinning a model. Therefore, sample sizes are defined in terms of numbers of observations (i.e. Table 2a) rather than numbers of individuals (Table 3a). Only sightings definitely identified to species are included so the sample size is the total number of observations across all species and all surveys but excluding those sightings not definitely identified to species: This yields a sample size of 2537 observations (Table 2a).

Fitting a global detection function across species has a number of advantages:

- It provides large sample sizes for fitting relationships with other covariates.
- Because only a single global model is being fitted rather than a separate model for each species it reduces the chances of the models capturing spurious relationships with covariates, which can occur due to sampling error particularly when a large number of models are fitted.
- It provides an approach which scales well with multiple species.

It should be noted that this approach assumes the relationships with other covariates are shared across species. This approach has the advantage of providing sufficiently large sample sizes to take these other covariates into account even for less common species, where sample sizes would be inadequate to fit reliable relationships. The disadvantage of the approach is that it only considers the element of the effects of these covariates that are common across species, and ignores any interaction with species (i.e. if the effect of survey conditions or cluster size on detectability varied between species). However if we were to analyse the data as separate species, sample sizes would generally be inadequate to fit

reliable relationships with these other covariates, and so we suggest the use of this approach is justified.

To ensure accurate estimates of the numbers of individuals (abundance) of each species on each survey our detection function model must successfully capture any variation in detectability between species, survey and cluster size. Therefore, we defined as our base model, which any alternative model must improve upon, a detection function including these three variables as covariates. With cluster size untransformed all models fitted including this covariate either crashed or failed to converge. Scaling cluster size so that it only varied between 0 and 1 did not resolve this problem. Therefore, throughout all analyses we have log transformed cluster size, which did resolve the problem, and treated it as a quantitative variable. Survey was treated as an 18 level categorical variable with one level for each survey date (Table 1b).

We compare the fit of alternative models using Akaike's Information Criterion, the best fitting model having the lowest AIC score. AIC is defined as

$$AIC = -2 \cdot \log_e(\lambda) + 2q$$

Where  $\log_e(\lambda)$  is the log-likelihood function evaluated at the maximum likelihood estimates of the model parameters and  $q$  is the number of estimated parameters in the model. We can interpret the first term as a measure of how well the model fits the data, while the second term is a penalty for the addition of parameters. The difference in fit from one model to another is measured as the change in the AIC value,  $\Delta AIC$ , with the better fitting model having the lowest AIC score. In the tables comparing the goodness of fit of different detection functions, we present  $\Delta AIC$  values comparing each model to the null model, with no covariates. We also present separately the changes in the first and second terms of the AIC score, so that the separate contribution of changes in the quality of fit and changes in the numbers of parameters to the AIC score can be assessed.

The null model with no covariates had an AIC score of 6903.9 (Table 6). When the three covariates included in the base model (i.e. species, log transformed cluster size and survey) were introduced singly into the model, all three led to a reduction in the AIC score compared to the null model, with species being the covariate which by itself provided the best fitting model. Reducing the AIC score by 107.4. For all three of the models in which the base model covariates were introduced in pairs, the AIC scores were reduced compared to either of the models including the same covariates singly. The best model including two covariates was the model including species and survey, which reduced the AIC score by 63.5 compared to the model just including species, the best fitting of the single covariate models. Adding log transformed cluster size into the best fitting of the two covariate models, which already included species and survey, further reduced the AIC score by 29.8. Thus, the inclusion of all three covariates improves the fit of the base model, and can be fully justified.

Survey date per se will not directly determine the probability of detection and the differences between surveys will reflect differences between surveys in environmental conditions (e.g. sea state, light levels) and other factors (e.g. observers). We considered the following covariates as alternative to the survey covariate in the base model to see if they could potentially explain the variation between surveys in detectability (Table 7):

- Sea state (the Douglas score as a quantitative variable).
- Wind force (the Beaufort score as a quantitative variable).
- Swell height (estimated in metres, as a quantitative variable).
- Observer (as a ten level categorical variable, one level for each observer).

We considered models in which the three environmental variables (sea state, wind force and swell height) were considered separately, and also models in which they were combined with the observer covariate. Initially, the environmental variables were only entered singly

into the models. The same survey vessel was used throughout and thus was not considered as a potential explanatory variable.

Only one of the models in which single environmental variables, sometimes in combination with the observer variable, were used as alternatives to the survey covariate provided a better fit to the data than the base model: this was the model in which sea state and observer were used as an alternative to the base model, which reduced the AIC score by 28.2 compared to the base model. In contrast, all of the models where the alternative explanatory variables were offered as supplements to survey provided a better fit to the data than the base model. However, none of these models provided as good a fit to the data as the best model identified previously, that where observer and sea state were used as alternatives to survey: thus this remained our best model. As one final test we investigated whether adding swell height to the current best model with sea state and observer provided a better fitting model; it did not. We did not consider introducing wind force into this model along with sea state as these two covariates are highly correlated with one another. Thus, our best model, which we will use to predict the probability of detection for common species is that which includes sea state and observer as explanatory variables along with species and log(cluster size.)

The best fitting model (sea state and observer instead of survey) had 7 less parameters than the base model, and thus provided a simpler, more parsimonious explanation of the data (AIC Term 2 equals 40 compared to 54). However, not only did it provide a simpler fit to the data, it also provides a closer fit (AIC Term 1 equals 6635.0 compared to 6649.3. Thus, the best model was both simpler ( $\Delta$  AIC Term 2 = -14) and provided a better fit to the data ( $\Delta$  AIC Term 1 = -14.2) than the base model. This suggests that these differences in observer and sea state could potentially explain all the variation between surveys and also explain additional variation within surveys.

The coefficients for Model 18 (Table 7) which are used as the detection function to predict the probabilities of detection for common species are presented in Table 8.

#### 4.1.1 Fit of the model to distance data

Scaled histograms of detection distances with the fitted detection function superimposed suggests that the model provides a reasonable fit to the data, across all observations (Figure 3a), for different species (Figure 4), cluster sizes (Figure 5), surveys (Figure 6), sea states (Figure 7) and observers (Figure 8).

A notable pattern in the data is the apparent under recording of birds within the innermost 0-50m distance band, with less birds being recorded in this distance band than in the 50-100m (Figure 3a). This pattern is present across species (Figure 4). There is also a clearer pattern for larger clusters to be recorded at greater distance from the vessel, and only for single birds does the density of birds recorded decline monotonically with increasing distance from the vessel (Figure 5). The pattern of more birds being recorded in the 50-100m distance band than in the innermost 0-50m distance band was recorded during 13 out of 18 surveys (72%), or 11 out of 14 surveys (77%) if we exclude surveys with less than 50 sightings (Figure 6). The pattern of more birds being recorded in the 50-100m distance band than in the innermost 0-50m distance band was recorded for 4 out of 6 sea states (67%), or 3 out of 4 sea states (75%) if we exclude sea states with less than 50 observations: There was no trend for under recording within the innermost distance bands to be associated with higher sea states (Figure 7). The pattern of more birds being recorded in the 50-100m distance band than in the innermost 0-50m distance band was recorded for 5 out of 10 observers (50%), or 3 out of 5 observers (60%) if we exclude observers with less than 50 observations (Figure 8).



Thus, the apparent under-recording of birds within the innermost distance band is not restricted to particular species, observers, surveys, sea states or cluster sizes. The underlying causes for this apparent under recording within the innermost distance band are unclear, but potentially could bias our estimates of the probabilities of detection. In particular, any under recording of birds within the innermost distance band will tend to lead to the probability of detection being overestimated, and thus density and abundance being underestimated.

#### 4.1.2 Variation in the probability of detection between species

Based upon the common species detection function model, for the 8 most common species, Table 9 presents estimates of the average probability of detection for both clusters and individuals across all surveys. Tables 15 and 16 present similar estimates but for each survey separately, with Table 15 presenting the probabilities of detection for clusters, and Table 16 presenting the probabilities of detection for individuals. The more birds a cluster holds the easier it is to detect (Table 10, Table 11). Therefore the probabilities of detection for individuals are generally higher than those for clusters (Table 9, and compare tables 15 and 16).

The average probability of detection for individuals from the common species detection function model for the 8 most common species (excluding the “Other” category) varies nearly two fold from 50% to 94% (Table 9, Figure 9). For 6 out of these 8 most common species, the probabilities of detection for individuals are less than would be obtained using the JNCC correction factors (Table 9 Stone *et al.* 1995), so that our estimates of density and abundance will tend to be higher. The exceptions are great skua, where the average probability of detection for individuals across surveys was estimated at 90% compared to the 77% that would be expected on the basis of JNCC correction factors and kittiwake, where the average probability of detection for individuals across surveys was estimated at 73% compared to the 71% that would be expected on the basis of JNCC correction factors. These are the two species with the lowest sample sizes (32 and 33 observations respectively), for which our estimates of the probability of detection are likely to be least reliable.

A limitation of relying on the average probabilities of detection to compare the probabilities of detection between species is that species may also differ with respect to other covariates that influence the probability of detection, preventing a clean comparison. For example, one species may tend to occur in larger clusters than another species, or, if recorded on different surveys at different times of year, tend to be recorded by different observers, or under different sea states. Thus, as well as presenting average probabilities of detection for each species we also present “standardised” probabilities of detection for each species (Table 9). These standardised probability of detection are for single birds in sea state 2 by observer 8. We use a single bird as this is the most frequently recorded cluster size, with 1529 out of 2537 observations (60%), Table 10). We use sea state 2 as this is the most frequently recorded sea state with 995 out of 2537 observations (39%), Table 12). We use Observer 8 as this is the observer with the most observations, with 606 out of 2537 observations (24%), Table 13). In the following sections we also compare standardised probabilities of detection between cluster sizes, sea states, observers and surveys. To standardise in these comparisons for species we use guillemot as our standard. We use guillemot as this is the most frequently recorded species with 895 out of 2537 observations (35%), Table 9).

The average probability of detection for individuals from the common species detection function model for the 8 most common species (excluding the “Other” category) varies across species by a factor of 1.9 from 50% to 94% (Table 9, Figure 9). In comparison, the standardised probability of detection (i.e single bird, sea state 2, observer 8) varies by a

factor of 2.7 from 34% to 92%. Thus, controlling for other sources of variation other than that with which we are principally concerned (i.e. variation between species) strengthens the apparent differences between species in the probability of detection. As would be expected the standardised probabilities of detection, which are for a single bird, are invariably lower than the average probabilities, which are for a range of cluster sizes. Most sightings are of single birds (Table 5a) and so for most species although the standardised probability of detection is lower than the average probability of detection, it is similar: for fulmar, gannet, shag, great skua and puffin the value of the standardised probability of detection is between 78% and 98% of the value of the standardised probability of detection.

However, for three species, kittiwake, common guillemot and razorbill, the standardised probability is much lower relative to the average probability than for these other species. For kittiwake, the standardised probability of detection is 34% compared to an average probability of detection of 73%. Thus the value of standardised probability of detection is 46% of the value of the average probability of detection. For guillemot, the standardised probability of detection is 42% compared to an average probability of detection of 63%. Thus the value of standardised probability of detection is 67% of the value of the average probability of detection. For razorbill, the standardised probability of detection is 34% compared to an average probability of detection of 53%. Thus the value of standardised probability of detection is 63% of the value of the average probability of detection. These large differences between the average and standardised probability of detection for some species can greatly change our assessment of the detectability of these species relative to others.

For example, when species are ranked in decreasing order of the average probability of detection, kittiwake is ranked 4<sup>th</sup> out of 8, above all the diving species. In contrast when species are ranked in a similar fashion on the basis of the standardised probability of detection for a single bird, kittiwakes are ranked 8<sup>th</sup> out of 8, below all the diving species. This difference is explained by the tendency for kittiwakes to occur in much large clusters than the other species. The average probability of detection does not correct for any such difference between species in cluster size, whereas the standardised probability of detection does. Kittiwakes were recorded on the water in groups of up to 400 birds, with a mean and median cluster sizes across all surveys of 34 and 8 birds respectively. In contrast, across the other species, the mean cluster size varied from 1.6 to 4.2, and the median cluster size was 1 for six out of the seven species, and 2 for the seven species.

As another example, if we compare the three species of auk on the basis of the average probability of detection, the ranking of species follows what might be expected on the basis of body length (i.e. guillemot, body length 40cm, 1<sup>st</sup> at 63.1%, razorbill, body length 38cm, 2<sup>nd</sup> at 53.3%, puffin, body length 28cm, 3<sup>rd</sup> at 50.1%). However, the ranking based on the standardised probability of detection for a single bird, (i.e. guillemot 1<sup>st</sup> at 42.2%, puffin 2<sup>nd</sup> at 39.3%, razorbill 3<sup>rd</sup> at 33.6%) suggest a very different pattern, with razorbills now less likely to be detected than the smaller puffin. Thus, these results suggest that for a given cluster size razorbills are less likely to be detected than puffins, a result which is obscured by differences in other covariates when comparing average probabilities of detection.

To conclude, when we standardise the probabilities of detection for cluster size, sea state and observer, differences between species which are obscured by differences in these other covariates, in particular cluster size, when comparing average probabilities of detection are revealed. This suggests that such standardisation is essential if we are to identify true patterns within the data.

#### 4.1.3 Variation in the probability of detection with cluster size

The average probability of detection is strongly related to cluster size (Figure 10), with the average probability of detectability varying by a factor of 1.6 between single birds and clusters of more than 10 birds (50% compared to 80%, Table 10). For our standard species, common guillemot, cluster sizes varied between 1 and 50 birds. Over this range of cluster sizes, the standardised probability of detection for a single guillemot (i.e. in sea state 2 for observer 8) varied by a factor of 1.8 from 42% to 76% (Table 11). Kittiwake showed the greatest range of cluster sizes of all species, with cluster sizes varying from 1 to 400. Therefore we also present standardised probabilities of detection for kittiwakes, over this range of cluster sizes. (Table 11), For kittiwakes, over this range of cluster sizes (i.e. 1 to 400) the probability of detection varies by a factor of 2.5 from 34% to 83%. For comparison, over the same range of cluster sizes (i.e. 1 to 400), the probability of detection for our standard species, common guillemot, would vary by a factor of 2.1 from 42.2% to 88.4%.

#### 4.1.4 Variation in the probability of detection between sea states

Camphuysen et al. (2004) recommends that all survey effort should be conducted in sea states of four or less. As discussed previously, this was achieved on most surveys, although 2.7% of coverage on the 27/03/2012, and 4.6% of coverage on the 18/04/2012 was conducted in sea state 5. Thus, in Table 12 we show the variation in the average and standardised probabilities of detection from sea state 0 to sea state 5, but in the discussion which follows we will focus on the variation which occurs up to sea state 4. For clusters, the average probability of detection declines consistently with increasing sea state from 79% at sea state 0 to 42% at sea state 4, variation by a factor of 1.9 (Table 12, Figure 11). For individuals, the average probability of detection declines with increasing sea state from 91% at sea state 0 to 57% at sea state 4, variation by a factor of 1.6 (Table 12). As well as declining by a smaller factor, the average probability of detection for individuals declines less consistently with increasing sea state than does the average for clusters, remaining effectively unchanged at 68-71% over sea states 1-3. These results are consistent with increasing sea state mainly reducing the probability of detection for smaller groups of birds, so that it has a more marked effect on the probability of detection of clusters than individuals. The standardised probability of detection for a single guillemot declines from 57% to 30% over sea states 1 to 4, variation by a factor of 1.9, similar to that observed for clusters.

#### 4.1.5 Variation in the probability of detection between observers

Between observers (Table 13, Figure 12), the average probability of detection for clusters varied by a factor of 1.72 from 41% to 70%, the average probability of detection for individuals varied by a factor of 1.66 from 47% to 78%, and the standardised probability of detection for a single guillemot varied by a factor of 1.60 from 39% to 62%. However observers fall into groups: Major observers, all of whom recorded over 300 observations, and minor observers all of whom recorded less than 50 observations. If we only consider the major observers, the average probability of detection for clusters varied by a factor of 1.44 from 46% to 66%, the average probability of detection for individuals varied by a factor of 1.25 from 61% to 76%, and the standardised probability of detection for a single guillemot varied by a factor of 1.47 from 42% to 62%. Thus, variation between observers was considerably less marked when we considered only the major observers, but still far from negligible.

#### 4.1.6 Variation in the probability of detection between surveys

The average probability of detection for individuals varies by a factor of 1.8 between surveys from 43% to 77% (Table 12, Figure 13). On some surveys few birds were recorded (e.g. just

9 sightings of birds on the water during the 11/12/2012 survey, Table 11) and some of the apparent variation in the probability of detection between surveys could potentially be due to sampling error. However, even if we restrict consideration to those surveys with more than 100 sightings, average probability of detection still varies by a factor of 1.7 from 46% to 77%. Thus, these results suggest that the variation between surveys in the probability of detection is likely to be genuine, and of a similar order of magnitude to that observed between species. Some of this variation between surveys could potentially be due to variation in species composition between surveys. However, even with species there are similar levels of variation (Table 16):

- For fulmar, the probability of detection for individuals varies between surveys by a factor of 1.4 from 60% to 86%;
- For gannet, the probability of detection for individuals varies between surveys by a factor of 1.1 from 87% to 97%;
- For shag, the probability of detection for individuals varies between surveys by a factor of 2.0 from 41% to 83%;
- For great skua, the probability of detection for individuals varies between surveys by a factor of 1.4 from 65% to 92%;
- For kittiwake, the probability of detection for individuals varies between surveys by a factor of 2.6 from 33% to 83%;
- For guillemot, the probability of detection for individuals varies between surveys by a factor of 1.8 from 38% to 70%;
- For razorbill, the probability of detection for individuals varies between surveys by a factor of 2.2 from 31% to 68%.
- For puffin, the probability of detection for individuals varies between surveys by a factor of 1.9 from 33% to 62%.

Thus these results suggest that the variation in the average probability of detectability between surveys within species is of a similar order of magnitude to the variation between species. If for example, we had used JNCC correction factors to correct our abundance estimates rather than estimates of the probability of detection based upon analysis of the survey data itself then not only would our estimates of the overall probability of detection for each species been very different to those obtained from site-specific data, but also the variation in detectability between surveys would also have been missed. Thus genuine variation in species abundance between surveys would have been conflated with variation in the probability of detection.

## 4.2 Detection function modelling for rare species

To estimate the probability of detection for rare species ( $n < 30$ ) we again fit a single detection function across all species, both rare and common. Other than species, this model has the same covariates as the common species detection function model (i.e. cluster size, sea state and observer). However species is replaced by covariates describing species traits that are chosen to explain, as simply as possible, the variation in detectability between species. The underlying assumption of this approach is that a rare species will have similar probabilities of detection to a common species with similar traits. In particular, the models reported here replace species by two covariates, one a measurement of body size and the other a categorical variable classifying species into two categories on the basis of behaviour:

- Surface/Aerial Feeders. (including gannets, gull, skuas, terns and petrels)
- Surface divers (within the current data set, just auks).

We consider two measurements of body size, body length and body length<sup>2</sup>. We would expect the width of the image on the retina of a bird sitting on the water to be directly proportional to its body length. Furthermore, to the extent that we can successfully approximate the area of this image as a rectangle (or trapezium, or triangle, or ellipse) with its height equal to a fixed (across species) proportion of its width, then the area of this image

should be directly proportional to body length<sup>2</sup>. Therefore, there are strong *a priori* reasons to expect the detectability of birds on the water to be positively related to body length or body length<sup>2</sup>. The justification for categorising birds based on their diving behaviour is that surface divers would be expected to have lower detectability than surface/aerial feeders, as when they are under the surface they are temporarily not available for detection. Values for these covariates for each species are provided in Table 9.

The advantage of estimating the probability of detection for rare species using this approach rather than using JNCC correction factors is that it provides an estimate for the probability of detection that is survey specific rather than generic. Furthermore, it allows the effect of other covariates including cluster size and environmental covariates on the probability of detection for rare species to be taken into account.

We compare the fit of models which are the same as the common species model but with the species covariate replaced by these covariates describing species traits in various combinations with one another, and with the null, base and common species detection function models (Table 17). We evaluate models in which the two covariates describing body size and the covariate describing behaviour are included separately, and models in which the two body size measurement are included in combination with the behaviour covariate.

All models including behaviour as a covariate provide a better fit to the data than the base model, with the best rare species model overall being the model which also includes the body length<sup>2</sup> covariate as a measurement of body size (Table 17). This best fitting model had an AIC score which was only marginally higher than that for the common species detection function model ( $\Delta$  AIC= +0.4). This is because although the rare species model provides a considerably poorer fit to the data than the common species detection function ( $\Delta$  AIC Term 1 = +12.4) it is a simpler model with 6 less parameters ( $\Delta$  AIC Term 2 = -12). Although it provides a poorer fit to the data than the common species detection function, it still provides a better fit than the base model ( $\Delta$  AIC Term 1 = -1.8). Furthermore, on the basis of the overall AIC score, other than the final model selected for the common species detection function model (model 18 in Table 7) none of the alternatives considered for the common species detection model (models 1 to 17 in Table 7) provided a better fit. Thus, the AIC scores suggest that the rare species model provide a reasonable fit to the data compared to alternatives, in particular the common species detection function.

The improvement in fit gained from using body length<sup>2</sup> as the measurement of body size rather than body length *per se* is marginal ( $\Delta$  AIC = -0.5), but carries no disadvantages in terms of adding to model complexity. Therefore we will use this model (Model 8 in Table 17) as the basis for predicting the probability of detection for „rare“ species.

With one exception, the coefficients for the covariates other than those capturing species traits in the rare species model (i.e. the cluster size, sea state and observer covariates) all take similar values (Table 18) to their counterparts in the common species model (Table 8). The one exception is the coefficient for observer 5 which in the common species detection function (Table 8) takes the value -0.07 but in the rare species detection function (Table 18) takes the value -0.22. However, observer 5 was only responsible for recording 32 out of 2537 observations (1.3%) This suggests that overall the covariates used in the rare species model to capture the variation between species in detectability are capturing the same underlying patterns of the variation in detectability between species as the species covariate in the common species model, so that the behaviour of the model with respect to other covariates is similar.

For the 8 common species considered by both the common species and rare species detection function models, Table 9 allows us to compare estimates for the probability of detection between the two models. The more similar the predictions from the two models,

the more confident we can be that rare species model successfully captures the variation between species in their probability of detection, at least for the common species for which a test is possible. For these eight common species, the predictions of the average probability of detection from the rare species model are closely correlated with those of the common species model for both clusters ( $r^2 = 64\%$ ) and individuals ( $r^2 = 75\%$ ). The predictions for the standardised probability of detection for a single bird are also closely correlated between the two models ( $r^2 = 70\%$ ). These strong correlations suggest that the rare species model successfully captures much of the variation between species in detectability, at least in relative terms.

For each species, expressing the average probability of detection for an individual bird from the common species model relative to its counterpart from the rare species model allows us to assess the average proportional change using the rare species detection function in place of the common species detection function is likely to cause in the resulting abundance and density estimates. This approach shows that for five out of the eight species (fulmar, puffin, common guillemot, great skua and gannet) the abundance and density estimates provided by the rare species detection function would on average be higher than those provided by the common species detection function: 0.2% higher for fulmar, 1.4% higher for puffin, 2.6% higher for guillemot, 9.7% higher for great skua, and 10.4% higher for gannet. For the other three species (shag, razorbill and kittiwake) the relative values of the average probability of detection for individuals from the two models suggests that the abundance and density estimates provided by the rare species detection function would on average be lower than those provided by the common species detection function: 6.9% lower for shag, 12.4% lower for razorbill and 19.0% lower for kittiwake.

A discrepancy between the predictions of the rare species model and those of the common species model could reflect either inaccuracies in the predictions of the rare species model, or inaccuracies in the predictions of the common species model, or inaccuracies in both. The predictions of the rare species model could be inaccurate because it is missing one or more of the covariates that actually determine the variation between species in the probability of detection. As the common species model explicitly includes species as a covariate, it cannot miss an underlying cause of variation between them in the same fashion. However, where sample sizes are small for a particular species, sampling error may mean that the common species model provides a less accurate prediction of the probability of detection for that species than the rare species model. The predictions from the rare species detection function model will also tend to be more accurate for common species than rarer species because where two species share similar values of the model's covariates, but one is much commoner than the other, the fitting of the model will be dominated by the distance data for the more common species.

As noted previously, although our definition of a common species being one with 30 or more observations is consistent with the guidance provided by Maclean *et al.* (2009) it is considerably less than the 60-80 observations recommended by Buckland *et al.* (2001:228) as the minimum required for reliable fitting of the detection function.

For the three species with sample sizes greater than 400 observations (guillemot, fulmar and puffin) the predictions of the two models are very similar to one another, with the rare species model expected to provide density and abundance estimates just 0.2 -2.6% greater than those provided by the common species model. This is to be expected both because with such large sample sizes, sampling error will only have a minimal impact on the predictions of the common species detection function model, and also because the fitting of the rare species model will tend to be dominated by the distance data for these three species (together they constitute 2121 out of 2578 sightings, 82%). In contrast, for the four species with sample sizes of less than 80 sightings (gannet, shag, great skua, kittiwake) the discrepancy varies from 6.9% to 19%, and for the three species with sample sizes of less

than 50 sightings (gannet, great skua, kittiwake) the discrepancy varies from 10.4% to 19.0%. With these small sample sizes, the discrepancies between the common species and rare species detection function models for these four species could potentially be explained by sampling error in the predictions from the common species detection function model, without having to invoke inaccuracies in the predictions from the rare species detection function. The one species where there is strong apparent evidence for the rare species detection function failing to capture the observed patterns in the data is razorbill, for which the density estimates as predicted by the rare species model are expected to be 12.4% lower than those from the common species model, and with a sample size of 197 negligible sampling error is expected in the predictions of the common species detection function. Both guillemots and razorbills are diving birds, with very similar body lengths (40 cm for guillemot, 38 cm for razorbill). Thus, the rare species detection function predicts a very similar probability of detection for these two species, with a probability of detection for individuals of 61.6% for guillemot and 60.9% for razorbill, a difference of 0.7%. In contrast the common species detection function model suggests a much greater difference between the probabilities of detection for the two species, 63.1% for guillemot, and 53.3% for razorbill, a difference of 9.8%. Thus, this suggests that the probability of detection for razorbill in this study was much lower than would be expected given its body length, and so was not fully captured by our rare species model.

In conclusion, these results suggest that the rare species detection function is likely to capture much, but not all, of the variation in the probability of detection between species. This conclusion is strengthened by the observation that the species for which the discrepancy between the predictions of the rare species detection function and the common species detection function is greatest are generally those for which the sample size of observations is less than 50, so that sampling error in the predictions of the common species detection function is likely to explain at least some of the discrepancy. However, even if we make the worst case assumption that the results of the common species detection function are completely accurate for all species, irrespective of sample size, and all the error lies within the rare species detection function, then the worst discrepancy we observe (for kittiwake) is an underestimate by 19% of abundance and density for the rare species model compared to the common species model.

For rare species with fewer than 30 sightings, the alternative to estimating the probability of detection using the rare species detection function model would have been to use the JNCC correction factors provided in Stone et al. (1995). Similar correction factors, sometimes taking into account variation between sea state, vessel type and source database, are presented in Skov et al. (1995). Here, we focus on the corrections provided by Stone et al. (1995) as these are the ones Maclean et al. (2009) recommend should be used. However, for Black guillemot, Stone et al. do not provide a correction factor and so for this species we use the estimate provided by Skov et al.

The average detection of probability for individuals determines the associated abundance and density estimates. For 11 of the 13 rare species for which the rare species detection function model will be used to provide abundance and density estimates the estimate of the average probability of detection for an individual as provided by the rare species detection function model is less than that implied by the JNCC correction factor for that species (Table 9). This implies that for these 11 species the abundance and density estimates we will obtain using the rare species detection function will tend to be higher than those we would have obtained if we had used the JNCC correction factors. Thus, any assessment of impacts is likely to be more conservative. The two exceptions are Arctic tern and great black-backed gull.

For great black-backed gull the difference between the average probability of detection for individuals from the two approaches is small: 71.4% from the JNCC correction factor

compared to 74.2% from the rare species detection function model. The 11 great black-backed gulls recorded were all single birds, so differences in cluster size cannot explain any difference in the observed probability of detection. However, the difference between the two estimates is easily explained by differences in survey conditions between this survey and those upon which the JNCC correction factor is based. Although using the rare species detection function to estimate the probability of detection for great black-backed gull will tend to yield slightly lower abundance estimates than would using the JNCC correction factors, the resulting estimates have the advantage of taking into account local survey conditions, including those such as observer and sea state that vary between surveys (Table 19 and 20).

For Arctic tern the estimates of the average probability of detection for individuals from the two approaches differ greatly from one another: an estimate of 82.5% from the rare species detection function model compared to 58.8% from the JNCC correction factor. However, in this study there was only one observation of Arctic terns on the water, and this was a group of 8 birds. The standardised probability of detection for a single bird is 58.9%, very similar to the probability of detection of 58.8% suggested by the JNCC correction factor. Thus, the relatively high average probability of detection for individual Arctic terns suggested by the rare species detection function can be explained by the fact that the estimate is for a single observation of 8 birds, for which, given the cluster size, irrespective of species a high probability of detection would be expected.

For 6 of the 8 common species the estimate of the average probability of detection for an individual as provided by the common species detection function model is less than that implied by the JNCC correction factor for that species (Table 9). This implies that for these 6 species the abundance and density estimates we will obtain using the common species detection function will tend to be higher than those we would have obtained if we had used the JNCC correction factors. Thus, any assessment of impacts is likely to be more conservative. The two exceptions are great skua and kittiwake. For kittiwake using the common species detection function to estimate the probability of detection rather than the JNCC corrections will on average lead to abundance and density estimates being 2% lower. For great skua using the common species detection function to estimate the probability of detection rather than the JNCC corrections will on average lead to abundance and density estimates being 14% lower.

Thus, for most species, using probability of detection estimates based upon either the common species detection function model, for the 8 common species, or the rare species detection function model, for the 13 rare species, will tend to yield higher abundance and density estimates than we would have obtained using JNCC correction factors. Thus, our assessment of impacts will tend to be more conservative than we would have been obtained using the standard JNCC correction factors. For four species, (great black-backed gull and Arctic tern from the rare species and great skua and kittiwake from the common species) our approach will tend to yield lower abundance and density estimates, and thus less conservative impact assessments than would using the JNCC correction factors. However, for three of the four species (great black-backed gull, great skua and kittiwake) the effects on the abundance and density estimates are small, whilst for the fourth species, Arctic tern, the results reflect the effect of cluster size. Furthermore, although in the interests of transparency we feel it is important to assess how the abundance and density estimates from our approach compare to those we would have obtained if we had used JNCC correction factors, we suggest our estimates of the probability detection based upon fitting detection functions to site specific data, and including covariates such as cluster size, sea state and observer are likely to provide a far more accurate estimate of the actual probability of detection than using generic JNCC correction factors. This includes capturing the considerable variation in the probability of detection between surveys for individual species



(Tables 15 and 16 for common species, Table 19b and 20b for rare species), which the use of a constant correction factor would ignore.

## 5 Estimating Density and Numbers

### 5.1 Overview

For each species, we calculated density and abundance estimates for each survey using the Horvitz-Thompson like estimator (Thomas *et al.* 2010, Borchers and Burnham 2004) provided by the *dht* (Density Horvitz-Thompson) function from the *mrds* (Mark Recapture Distance Sampling) *R* package (Laake *et al.* 2012). We did this for birds in flight and birds on the water separately. For birds on the water, for species with 30 or more sightings the common species detection function model is used to estimate the probability of detection. Otherwise the rare species detection function model is used. For birds in flight all birds in snapshot are assumed to be detected.

The variance of the density and abundance estimates provided by *mrds* has two components: 1) uncertainty in the estimate of the probability of detection and 2) uncertainty in the encounter rate estimate. With respect to estimating the contribution of the variance in encounter rate to the overall variance estimate, we use the default option (*varflag*=2). For birds on the water, the detection function models fitted using *mrds*'s *ddf* function provide estimates of the variance associated with the probability of detection estimates which are used by the *dht* function to estimate the contribution from this source to the overall variance estimate. For birds in flight we fit our own customised detection function model, which assumes 100% probability of detection for all sightings, with no variance in the estimates. Using this customised detection function model allows us to use *dht* to estimate density and abundance for birds in flight, providing confidence limits which take into variance in encounter rate.

This approach provides abundance and density estimates with associated variance estimates, and thus confidence limits, for both birds on the water and birds in flight. The variance estimate for birds on the water includes contributions from both uncertainty in the estimate of the probability of detection and uncertainty in the estimate of the encounter rate. For birds in flight, only variance in encounter rate contributes to this variance estimate, with zero contribution assumed from uncertainty in the estimated probability of detection.

Having calculating separate density and abundance estimates for birds on the water and birds in flight we then combine these estimates to provide an overall estimate, with associated confidence limits for birds on the water and birds in flight combined (See Appendix A).

Using the above approach we calculate density estimates for each survey. From these density estimates for each survey we derive density estimates for species specific seasons, as described below.

We have used 90% two sided confidence limits, which correspond to 95% one sided confidence limits throughout. Thus there is an estimated 90% probability that the true estimate lies between the lower and upper confidence limits, and an estimated 95% probability that the true estimate lies either below the upper confidence, or above the lower confidence limit. For the purposes of site characterisation, we are often interested in determining whether the number of birds occurring within a given area falls below some threshold value which defines a species of conservation interest. Thus, the upper confidence limit is generally of more interest than the lower confidence limit. Furthermore we are generally only interested in the probability the true estimate falls below this upper confidence

limit, irrespective of whether or not it falls above some lower confidence limit: Thus we are primarily interested in one tailed rather than two-tailed confidence limits, and this is why we have chosen to use confidence limits corresponding to 95% one side confidence limits, rather than two sided confidence limits.

In these analyses no distinction has been made between different age classes of birds. Thus the abundance and density estimates presented are for all age classes of birds combined. Where the impact assessment requires abundance or density estimates for a particular age class, these will be derived by multiplying these abundance estimates by the estimated proportion of birds within that age class from a separately derived age class distribution.

Of the 2573 sightings of birds on the water, 36 were not identified to species. Four of these sightings (across three surveys) were recorded as "Auk sp." and 32 (across 6 surveys) were recorded as "guillemot or razorbill". Of the 1861 sightings of birds in flight, 16 were not identified to species. Fifteen of these sightings (across 8 surveys) were recorded as "guillemot or razorbill" and a single sighting was recorded as "large gull sp."

For the species concerned, as well as presenting uncorrected abundance estimates, we also present abundance estimates corrected for the presence of these sightings not identified to species. In particular, we correct the estimates for recorded auk species (i.e. guillemot, razorbill, puffin, black guillemot and little auk here) for the presence of birds recorded as "guillemot or razorbill" and "auk sp" and we correct the estimates of large gulls (i.e. herring and great black-backed gull here) for the presence of birds recorded as "large gull sp.". The procedure we use to make these corrections is described fully in Appendix C.

Across all tables:

- Densities are in units of nos/km<sup>2</sup>.
- "est." is an abbreviation of "estimate".
- "cv" is an abbreviation of "coefficient of variation".
- "df" is an abbreviation of "degrees of freedom".
- "lcl" is an abbreviation of "lower confidence limit".
- "ucl" is an abbreviation of "upper confidence limit".
- "Max. est" is an abbreviation of "maximum estimate", and
- "Max ucl" is an abbreviation of "maximum upper confidence limit".

Where "Corr." appears in front of a species name, these density and abundance estimates have been corrected for the presence of sightings not identified to species. In calculating these corrected estimates, we also need to calculate abundance estimates for relevant taxonomic grouping above the species level(Appendix C). In the tables provided:

- "All auks" identifies estimates for all auks combined, including those not identified to species.
- "PI auks" identifies estimates for all auks combined, based solely on sightings identified to species.
- "All G and R" identifies estimates for guillemot and razorbill combined, including sightings not identified to species.
- "PI G and R" identifies estimates for guillemot and razorbill combined, based solely on sightings identified to species.
- "All large gulls" identifies estimates for all large gulls combined, including those not identified to species.
- "PI large gulls" identifies estimates for all large gulls combined, based solely on sightings identified to species.

## 5.2 Seasonal abundance estimates

From the density estimates for each survey, for species with more than 5 records (in flight and on the water combined) we derive abundance estimates for species-specific seasons with associated confidence limits using the approach set out in Appendix B. The confidence limits are based upon bootstrapping and take into both the uncertainty within the estimates of density for each survey and also the uncertainty resulting from the limited number of surveys within each season providing a sample of the actual temporal variation in density within a season. Table 21 summarises the seasons used to characterise each species. The justification for the seasons used is provided in the main technical report (REFERENCE). We present two sets of seasonal abundance estimates for each species:

- Estimates across years, with a single estimate for each season across the years.
- Estimates by years, with separate estimates for each season in each year.

For the estimates by year, where a season extends over two calendar years, it is identified by the year it starts in. For example, a “Winter 2012” season would begin in the autumn of 2012, and finish in the spring of 2013.

Survey work began in late March 2012 and finished in early March 2014. For species with breeding seasons that start after mid March, the initial surveys in early 2012 are within their 2011 winter season. However, given the limited coverage confined to the end of the season, year specific estimates for winter 2011 would be unreliable and unrepresentative. Furthermore, if we were to include surveys that fall within the winter 2011 season for a species in the calculation of seasonal estimates across years for that species, then with complete coverage of the 2012 and 2013 winters, including surveys completed late in the winter of 2011 would bias the sample towards late winter. Therefore, to provide a more representative sample when calculating seasonal abundance estimates for species with breeding seasons beginning after mid-March we have excluded any surveys conducted in early 2012 that fall into the winter 2011 season for that species. Similarly, for species with breeding seasons that begin before mid-March, the final surveys in early 2014 provided limited coverage of the start of their 2014 breeding season. Again, so as to provide a more representative sample on which to base our seasonal abundance estimates we have excluded any such surveys when calculating the seasonal abundance estimates for that species.

Thus, all our seasonal abundance estimates, whether by or across years, are based on the 2012 and 2013 breeding and non-breeding seasons, for which for complete, or near complete coverage, was achieved.

In March 2013, in addition to complete surveys conducted on the 5<sup>th</sup> March and 30<sup>th</sup> March, the latter in lieu of the April survey, an incomplete survey was undertaken on the 4<sup>th</sup> March. When presenting results for individual surveys, the abundance estimates presented for individual surveys include estimates for this partial survey. However when calculating seasonal abundance estimates we have excluded this partial survey on the 4<sup>th</sup>, so as to provide a more representative sample.

The original intention was to follow standard practice and base our impact assessment on monthly abundance estimates derived from the single survey of 11 transects on a single day conducted in each month. However, with a single survey within a month, there is no way to distinguish within month variation from between month variation in abundance. Therefore instead we decided to base our impact assessment on seasonal abundance estimates, with each season generally consisting of several months, and thus with density estimates based on several surveys. This allows between and within season variation to be distinguished. As seasonal estimates are based on several surveys, they should be more reliable, and less prone to sampling error, than would monthly estimates, based upon a single survey. These

seasonal abundance estimates are calculated on the assumption that the surveys conducted within each season provide a representative sample, and also that the estimates of density and abundance for each survey date can be assumed to be statistically independent.

As well as presenting estimates for the average density or abundance of birds present across each season we also present estimates for the maximum density/abundance of birds potentially present, all with 95% confidence intervals.

### **5.3 Spatial strata for which we estimate density and abundance**

To assess the effect of a particular impact on a particular species we require density and abundance estimates for an appropriate Anticipated Impact Footprint (AIF). The size of the appropriate AIF can vary according to species and the particular impact under consideration. Therefore, in order to provide flexibility we provide abundance and density estimates for the following strata (Figure 14, Table 22):

1. The Development Area
2. The Development area with a 1 km buffer.
3. The Development area with a 2 km buffer.
4. The Development area with a 3 km buffer.
5. The whole survey area.

Table 22 provides details of each of these strata, including the survey effort and numbers of transects that would be expected to be covered if full coverage is achieved during a single survey.

### **5.4 Overview of presentation of density and abundance estimates**

The raw results of the analyses estimating density and abundance are presented in the following appendices:

- Appendix D details the survey effort actually achieved on each survey date in each strata.
- Appendix E details the density estimates for each species, on each survey for each strata, based on survey effort within that strata. Separate density estimates are provided for birds on the water, birds in flight, and both of these combined.
- Appendix F presents density estimates for species specific seasons calculated across years, for each species in each strata, based on survey effort within that strata.
- Appendix G also presents density estimates for species specific seasons for each species in each strata, based on survey effort within that strata, but this time calculated within years, whereas the estimates presented in Appendix F are calculated across years.

### **5.5 Temporal variation in density estimates**

For each species, Figures 1 to 24 in Appendix H plots abundance estimates versus date. The purpose of these figures is to reveal patterns of temporal variation in the abundance of species. The abundance estimates displayed are for the whole survey area for birds on the water and birds in flight combined. Both of the figures presented for each species display the abundance estimates for each survey, with seasonal abundance estimates superimposed. The difference between the two figures is that whereas the seasonal abundances shown in the left hand figure are calculated across years, those displayed in the right hand figure are calculated within years. All abundance estimates are accompanied by 95% one sided (90% double sided) confidence limits.

One common pattern which occurs across species is the large variation in abundance estimates apparent between survey dates within a season. Particularly given the limited number of surveys usually available for calculating abundance estimates for each season this means that the accuracy of the seasonal abundance estimates and their associated confidence limits will be heavily dependent on the assumption that the surveys provide a representative sample of the actual temporal variation in density occurring within a season. However, seasonal abundance estimates are still likely to provide a more reliable basis for assessing impacts than either monthly estimates or individual survey estimates as the same difficulties with temporal variability would also apply to these estimates, but the sample sizes available in terms of number of surveys to try and assess this variability are smaller.

The confidence limits for the seasonal abundance estimates are for the average of the individual abundance estimates across the whole season rather than for the individual estimates themselves. This is why they are generally much narrower, and more tightly defined, than the confidence limits associated with the estimates for the individual surveys.

For the following species sample sizes were insufficient to allow the calculation of seasonal abundance estimates, or to formally consider differences in abundance and density estimates between seasons and years:

- Small numbers of red-throated divers (i.e. single birds on each of three surveys, 2 in flight, 1 on the water) were recorded prior to the breeding season (i.e. in February or March) in 2012 and 2014 but not 2013 (Figure H.1).
- A single black-throated diver was recorded on the water in March 2014, prior to the breeding season (Figure H.2).
- A single great northern diver was recorded on the water in June 2013 (Figure H.3).
- A single sooty shearwater was recorded in flight on autumn passage in September 2013 (Figure H.5).
- There was a single record of 2 cormorants on the water in October 2013 (Figure H.9).
- A single black-headed gull was recorded in flight in June 2013 (Figure H.13).
- There was a single record of two common terns in flight in June 2013 (Figure H.18).
- A single little auk was recorded in flight in March 2013 and two birds, one in flight and one on the water, were recorded in February 2014 (Figure H.23).

Summer visitors recorded only during the breeding season in both years were Manx shearwater (Figure H.6), storm petrel (Figure H.7), Arctic skua (Figure H.11), great skua (Figure H.12) and Arctic tern (Figure H.19). A single great skua recorded on the 27<sup>th</sup> March 2012, was just outside the formal April-September breeding season defined for this species.

Species consistently more abundant in the breeding season than in the winter season in both years were kittiwake (Figure H.17), common guillemot (Figure H.20), razorbill (Figure H.21) and puffin (Figure H.24).

Species consistently more abundant during the winter months than in the breeding season in both years were herring gull (Figure H.15) and great black-backed gull (Figure H.16). Indeed, although herring gulls breed locally, birds were only recorded outside the breeding season, during the winter months (Figure H.15).

Species for which there was no clear consistent difference between the breeding season and winter season across both years were fulmar (Figure H.4), gannet (Figure H.6), shag (Figure H.10), common gull (Figure H.14) and black guillemot (Figure H.22). Of these species, for fulmar (Figure H.4), shag (Figure H.10) and black guillemot (Figure H.22), not only was there no clear consistent difference between the seasons across years, but abundance/density appeared to be relatively constant (compared to other species)

throughout the year. For black guillemot and shag this is consistent with locally breeding birds remaining close to their breeding colonies throughout the year (Wernham et al. 2002). Fulmars will also attend their colonies throughout most of the year (Wernham et al. 2002), and the relative stability (compared to other species) of the numbers of this species throughout the year might also reflect the presence of locally breeding birds outside the breeding season.

The month of August was defined as a distinct “chicks at sea/ Moulting” season for both guillemots and razorbills. A single survey was conducted during August 2012 (on the 20th) and no surveys were conducted during August 2013. During the single survey in August 2012, 7 razorbills were recorded, and no guillemots. Thus for guillemot the abundance estimate for this season is zero (Figure H.20), whilst for razorbill it is nonzero, but based on a single survey in August 2012 (Figure H.21). Of the 7 razorbills recorded, 3 were chicks, 1 was an adult in summer plumage, and 3 were adults in transitional plumage. This composition is consistent with a high proportion of the razorbills present in the survey area at the time of the survey being either chicks or adults accompanying chicks and undergoing moult.

For a given species in a given season, when seasonal abundance estimates are calculated by year the confidence limits for the abundance estimate for the first year generally show considerable overlap with the confidence limits for the abundance estimate in the same year, and a single seasonal abundance estimate calculated across years appears to provide a reasonable summary of the data across both years. Thus, there is little or no evidence to suggest that the accuracy of our impact assessments might be improved by considering years separately rather than calculating seasonal abundance estimates across years. This is to be expected given the limited sample sizes (in terms of number of surveys) available for calculating seasonal abundance estimates, particularly when these estimates are calculated within years. Thus, we will base our impact assessment on seasonal abundance estimates calculated across years.

Two cases where there are apparent differences between years in seasonal abundance are common gulls during the winter (Figure H.14) and great black-backed gulls during the breeding season (Figure H.16). Common gulls were recorded during the 2012 winter but not the 2013 winter whilst great black-backed gulls were recorded during the 2012 breeding season but not the 2013 breeding season. However, although these differences could potentially represent genuine differences between the seasons in the presence or absence of birds, they could also represent chance effects. Common gulls were recorded in 3 out of 5 surveys during the 2012 winter and during 0 out of 4 surveys during the 2013 winter. Thus, over the two winters common gulls were recorded during 3 out of 9 surveys. Thus, the overall probability common gulls were not recorded during a winter survey is  $6/9$ . Therefore, the probability they would not be recorded during any of the four surveys in the 2013 winter is  $(6/9)^4 = 20\%$ . Similarly, great black-backed gulls were recorded in 3 out of 5 surveys during the 2012 breeding season and during 0 out of 3 surveys during the 2013 breeding season. Thus, over the two breeding seasons great black-backed gulls were recorded during 3 out of 8 surveys. Thus, the overall probability black-backed gulls were not recorded during a breeding season survey is  $5/8$ . Therefore, the probability they would not be recorded during any of the three surveys in the 2013 winter is  $(5/8)^3 = 24\%$ . Therefore with probabilities of 20% and 24% in both cases the complete absence of records in one year is not the most likely outcome (i.e. the probabilities are both less than 50%) but is still likely to be a relatively common occurrence, without having to invoke differences between the seasons (i.e. the probabilities are both considerably greater than 5%). Thus, even for these two species, in these two seasons, we suggest it is reasonable to use densities calculated across seasons as the basis of our impact assessment.

## 6 References

Bivand, R. S., Pebesma, E. J., and Gómez-Rubio, V. 2008. Applied Spatial Data Analysis with R(Use R). Springer, 1st edition (August 2008).

Borchers, D.L. & Burnham, K.P. 2004. General formulation for distance sampling. Advanced Distance Sampling (eds S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers & L. Thomas), pp. 6–30. Oxford University Press, Oxford.

Buckland, S. T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press.

Buckland, S. T., Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. Advanced distance sampling. Oxford University Press.

Camphuysen, C.J., Fox, T., Leopold, M.F. & Petersen, I.K. 2004. Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the UK. A report for COWRIE.

Laake, J., Borchers, D., Thomas, L., Miller, D. & Bishop, J. 2012. mrds: Mark-Recapture Distance Sampling. R manual. R package version 2.1.0. URL: <http://CRAN.R-project.org/package=mrds>.

Maclean, I.M.D, Wright, L.J., Showler, D.A. & Rehfisch, M.M. 2009. A Review of Assessment Methodologies for Offshore Windfarms. British Trust for Ornithology Report Commissioned by COWRIE

Milliken, G.A. & Johnson, D.E. 2009. Analysis of Messy Data Volume 1: Designed Experiments, Second Edition. 674 pages. CRC Press. ISBN 9781420010152.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Satterthwaite, F. E. 1946. "An Approximate Distribution of Estimates of Variance Components.", *Biometrics Bulletin* 2: 110–114, [doi:10.2307/3002019](https://doi.org/10.2307/3002019)

Stone, C.J., Webb A., Barton C., Ratcliffe N., Reed T.C., Tasker M.L., Camphuysen C.J. & Pienkowski M.W. 1995. An atlas of seabird distribution in north-west European waters. JNCC, Peterborough.

Thomas, L., S.T. Buckland, E.A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R.B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* 47:5-14. DOI: 10.1111/j.1365-2664.2009.01737.x

Wernham, C.V., Toms, M.P., Marchant, J.H., Clark, J.A., Siriwardena G.M., & Baillie S.R. (eds). 2002. The Migration Atlas: movements of the birds of Britain and Ireland. T. & A.D. Poyser, London.

## Appendix A: Combining density estimates for birds on the water and birds in flight

For data collected under the standard ESAS methodology (Camphuysen *et al.* 2004), as in this study, distance sampling methods can only be applied to birds recorded on the water within transects, as for birds in flight no distance data are available. Thus, for birds in flight we have assumed that all birds within the transect were detected. However, the data for birds in flight was also analysed using the code from Distance, assuming 100% detection within the transect. This provides density estimates for birds in flight along with associated estimates of variance calculated within the same framework/software as the estimates for birds on the water. Whereas the variance estimate for birds in flight only includes components associated with the estimation of encounter rate and mean cluster size, the variance estimate for birds on the water also includes components associated with the estimation of the detection function.

Having obtained  $\hat{D}_w$ , the estimated density of birds on the water (with associated standard error  $s_w$  and degrees of freedom  $v_w$ ) and  $\hat{D}_f$ , the estimated density of birds in flight water (with associated standard error  $s_f$  and degrees of freedom  $v_f$ ) from separate Distance analyses, we compute the estimated density of all birds,  $\hat{D}_t$  as:

$$\hat{D}_t = \hat{D}_w + \hat{D}_f \quad (1)$$

Assuming  $\hat{D}_w$  and  $\hat{D}_f$  are uncorrelated random variables we estimate  $s_t$  the standard error of their sum  $\hat{D}_t$  as:

$$s_t = \sqrt{s_w^2 + s_f^2} \quad (2)$$

Applying the Satterthwaite Approximation for degrees of freedom (Satterthwaite 1946, Milliken and Johnson section 2.7, p.33) provides the following estimate for the degrees of freedom,  $v_t$ , associated with this estimate:

$$v_t = \frac{s_t^4}{\frac{s_w^4}{v_w} + \frac{s_f^4}{v_f}} \quad (3)$$

Using these estimates of the standard error,  $s_t$  and degrees of freedom,  $v_t$ , associated with the combined density estimate,  $\hat{D}_t$ , we derive confidence limits using equations 3.72 to 3.74 in Buckland *et al.* (2001:77), which assume that  $\hat{D}_t$  is log-normally distributed.



## Appendix B: Calculating seasonal abundance estimates

Assume we are deriving an average density estimate for a season in which  $n$  surveys took place. Let  $d_i$  be the density estimate for survey  $i$  with coefficient of variation  $c_i$  and degrees of freedom  $v_i$ . If we assume the surveys provide a representative sample of the variation in the true densities of birds present during the season, the best estimate for the average density of birds present over the season,  $D$ , is given by:

$$D = \frac{\sum_{i=1}^n d_i}{n} \quad (1)$$

There are two sources of uncertainty associated with this estimate of the average density of birds present over the season:

1. Uncertainty within the individual density estimates for each survey, reflecting the accuracy with which they capture the actual density of birds present at the time of survey.
2. Uncertainty because the limited number of surveys conducted within a season only samples the actual temporal variation in density through the season, leading to sampling error.

With the aim of capturing both of these sources of error we bootstrap confidence limits based on 1000 simulations for each species. For each simulation we choose a random set of surveys from the original surveys. This is achieved by randomly sampling (with replacement) a list of unique identifiers for each of the surveys. This yields a random sample of surveys with the same number of surveys as the original data set. For example, if a seasonal density estimate is based upon 6 surveys, numbered 1-6, the random sample upon which a particular simulation is to be based might consist of surveys 5,1,3,2,5 and 1. This first stage of the simulation aims to capture the error resulting from the limited number of surveys only providing a sample of the underlying temporal variation in the actual density of birds present. It assumes that the surveys provide a fully representative sample of the underlying temporal variation in density through the season, and uses the survey data to define the underlying distribution of densities from which our bootstrap samples are chosen. Although the distribution of densities across different dates within the survey data can be thought of as the most likely instance of the underlying distribution, the true underlying distribution will usually differ from that recorded, particularly given the limited number of surveys. The question we are answering is if we assume that we can accurately infer the shape of the underlying distribution of densities across different dates/times within the season from our sample of survey data, what are the consequences of the limited number of surveys.

For each survey  $j$  within a bootstrap sample  $k$  the distance sampling analyses provide a central density estimate  $d_{jk}$  with coefficient of variation  $c_{jk}$  and degrees of freedom  $v_{jk}$ . These define the expected distribution of the true value of the density for that survey around the central density estimate. This distribution is assumed to be a log normal distribution. For each simulation  $k$ , for each survey  $j$  we chose a density estimate,  $r_{jk}$  at random from this distribution, calculated as follows:

$$r_{jk} = d_{jk} \exp\left(t_{v_{jk}} \sqrt{\log_e(1 + c_{jk}^2)}\right) \quad (2)$$

Where  $t_v$  is a random number chosen from a Student's  $t$  distribution with  $v$  degrees of freedom. This is consistent with the approach taken to defining the confidence limits for density estimates from distance sampling analyses as set out in equations 3.72 to 3.74 of Buckland *et al.* (2001:77). In this second step of the simulation process we are capturing the potential error resulting from our estimate of the density of birds based on distance sampling

for a particular survey not accurately capturing the actual number of birds present at the time of that survey.

For each bootstrap sample  $k$  we calculate  $R_k$  the average density of birds across surveys, based upon the random sampled density estimates for that simulation:

$$R_k = \frac{\sum_{j=1}^n r_{jk}}{n} \quad (3)$$

We calculate the lower confidence limit (lcl) for  $D$ , the estimated average density across the season, as the 5% quantile of the  $R_k$  estimates over the 1000 simulations. We calculate the upper confidence limit (ucl) for  $D$ , as the 95% quantile of the  $R_k$  estimates over the 1000 simulations. This approach provides 95% one sided and 90% double side confidence intervals.

We also calculate the mean and median values of the  $R_k$  estimates over the 1000 simulations. The distance sampling software providing the density estimates for each of the surveys assumes the density estimates for a particular survey follows a log normal distribution. For a log normal distribution the central estimate corresponds to the median of the distribution. As a result of skewing, the mean of a log normal distribution tends to be greater than this central estimate and depends upon the variance of the distribution as well as its central estimate. The median of the  $R_k$  estimates tends to be close to our best/central estimate for the average density of birds present over the season,  $D$ , whereas the mean value tends to be greater, sometimes much greater than this central estimate. This suggests that the estimates for the average density of birds present over a season may, like the estimates for each survey, also follow a log normal distribution, with our best estimate representing the median of this distribution.

## Appendix C: Uncertain identifications

### Introduction

For some broader taxonomic groups consisting of similar species, it is not always possible to positively identify sightings to individual species although they can be assigned to a broader taxonomic grouping. For example, within the auk taxon, as well as sightings positively identified positively to species we have “uncertain” sightings potentially recorded as one of:

- Common guillemot or razorbill.
- Puffin or little auk.
- Auk sp.

For each species, density estimates corrected for uncertain identifications are calculated separately for birds on the water and birds in flight. These are combined to provide density estimates for all birds using the same approach used for all other species as set out in Appendix A.

For birds on the water, we cannot fit a detection function to the distance data for an “uncertain” category by itself as the numbers of sightings not identified to species will often tend to increase with increasing distance from the vessel as birds will tend to more difficult to identify the further they are from the observer. This violates a core assumption of distance sampling, that the items for which we are estimating density must be equally likely to occur at all distances from the observer. Therefore, instead of estimating the probability of detection and density for the uncertain identification category by itself we estimate the probability of detection and density for the whole taxon based upon all sightings including those not identified to species. We then compare this density estimate to that based solely on sightings that were successfully identified to species to estimate the proportion of birds present within the taxon that were not identified to species. We then use this proportion to correct the density estimates for uncertain identification.

For birds in flight, we assume a probability of detection of 100% rather than estimate a probability of detection by fitting a detection function to the distance data. R Therefore, we could potentially estimate the density of the “uncertain” category by itself without violating any key underlying assumptions. However, there is no disadvantage to applying the same approach as is required for birds on the water. Thus, we use the same approach for both birds on the water and birds in flight. Consider a taxonomic group for which sightings are either positively identified as one of  $n$  species or for which the species is recorded as uncertain. We calculate a corrected density  $c_i$  for species  $i$  as:

$$c_i = d_i \frac{D_U}{D_P} \quad (1)$$

Where:

- $d_i$  is the uncorrected density estimate for individual  $i$  based solely on positively identified individuals.
- $D_P$  is the density estimate for the whole taxon based only on positively identified individuals.
- $D_U$  is the density estimate for the whole taxon based on all sightings, including those that could not be identified to species.

To estimate the standard error,  $S_i$ , of this corrected density estimate as a first approximation we assume the components are independent and use the Delta method (Seber 1982:7-9):

$$S_i = c_i \sqrt{\left[ \frac{s_i^2}{d_i^2} + \frac{S_U^2}{D_U^2} + \frac{S_P^2}{D_P^2} \right]} \quad (2)$$

Where

- $s_i$  is the standard error estimate for  $d_i$ , the uncorrected density estimate for individual species  $i$ .
- $S_p$  is the standard error estimate for  $D_p$ , the density estimate for whole taxon based only on positively identified individuals.
- $S_u$  is the standard error estimate for  $D_u$ , the density estimate for the whole taxon based on all sightings, including those that could not be identified to species.

Applying the Satterthwaite Approximation for degrees of freedom (Satterthwaite 1946, Milliken and Johnson section 2.7, p.33) provides the following estimate for the degrees of freedom,  $V_i$ , associated with  $c_i$ , the corrected density estimate:

$$V_i = \frac{S_i^4}{\frac{S_i^4}{v_i} + \frac{S_p^4}{V_p} + \frac{S_u^4}{V_u}} \quad (3)$$

Where:

- $v_i$  is the degrees of freedom associated with  $d_i$ , the uncorrected density estimate for individual species  $i$ .
- $V_p$  is the degrees of freedom associated with  $D_p$ , the density estimate for whole taxon based only on positively identified individuals.
- $V_u$  is the degrees of freedom associated with  $D_u$ , the density estimate for the whole taxon based on all sightings, including those that could not be identified to species.

Using these estimates of the standard error,  $S_i$  and degrees of freedom,  $V_i$ , associated with the corrected density estimate for species  $i$ ,  $c_i$ , we derive confidence limits using equations 3.72 to 3.74 in Buckland *et al.* (2001:77), which assume that  $c_i$  is log-normally distributed.

Thus, given a density estimate for species  $i$ ,  $d_i$ , a density estimate for the whole taxon based solely on positively identified individuals,  $D_p$ , and a density for the whole taxon based on all individuals recorded, including those not identified to species ( $D_u$ ) we can derive a density estimate for species  $i$  corrected for uncertain identifications. We now describe how we obtain these three different density estimates.

#### **Obtaining estimates for $d_i$ , the density estimate for species $i$**

The uncorrected density estimates for individual species,  $d_i$ , along with their associated standard errors ( $s_i$ ) and degrees of freedom ( $v_i$ ) are the results of the distance sampling analyses described fully in the main report. The probabilities of detection used in calculating these estimates are calculated by fitting detection functions across multiple species. For birds on the water, the probability of detection for common species (with 30 or more sightings) is derived from the Common Species Detection Function (Table 9) and the probability of detection for rare species (less than 30 sightings) is derived from the Rare Species Detection Function (Table 18). For birds in flight, a customising detection function model implementing 100% detection is implemented.

#### **Obtaining estimates for $D_p$ , the density estimate for the whole taxon, based on just positively identified individuals**

The data used to estimate the uncorrected densities of individual species as described in the previous paragraph consists solely of sightings positively identified to species. Therefore, we estimate the density estimate for the whole taxon based solely on positively identified individuals  $D_p$ , by simply summing these density estimates for individual species:

$$D_p = \sum_{j=1}^n d_j \quad (4)$$

Where  $d_j$  is the density estimate for the  $j$ th species out of  $n$  species. Under the simplifying assumption that the density estimates for the individual species are independent of one another we estimate the standard error for their sum,  $S_p$ , as:

$$S_p = \sqrt{\sum_{j=1}^n s_j^2} \quad (5)$$

Where  $s_j$  is the standard error estimate for the  $j$ th species out of  $n$  species within the taxon. Applying the Satterthwaite Approximation for degrees of freedom (Satterthwaite 1946, Milliken and Johnson section 2.7, p.33) provides the following estimate for the degrees of freedom,  $V_p$ , associated with this estimate:

$$V_p = \frac{S_p^4}{\sum_{j=1}^n \frac{s_j^4}{v_j}} \quad (6)$$

Where  $v_j$  is the estimated degrees of freedom associate with the density estimate for the  $j$ th species out of  $n$  species within the taxon.

### **Obtaining estimates for $D_u$ , the density estimate for the whole taxon, including sightings not identified to species**

For a study where separate detection functions are fitted for individual species, estimating  $D_u$  is straightforward, as it simply entails providing the appropriate subset of observations (i.e. all observations within the taxon, including those not identified to species) to the ddf fitting the detection function and to the dht function estimating the density of birds using these probability of detection estimates. For this study, where single detection functions are fitted across species, a more complex approach is required, although the underlying principles remain the same.

For birds in flight, the customised ddf object implementing a 100% probability of detection for all sightings across species includes sightings that were not identified to species. Therefore, for birds in flight obtaining the density estimate for a whole taxon including uncertain identifications,  $D_u$ , is simply a providing matter of providing the dht function which computes density and abundance estimates with this ddf object along with the appropriate subset of observations (i.e. all sightings of birds in flight recorded in snapshot within the target taxon including sightings not identified to species) as the basis of its calculations. As well as directly providing estimates of  $D_u$ , the dht function also provides estimates of the standard error  $S_u$  and degrees of freedom  $V_u$  associated with this estimate.

For birds on water a more complex approach is required as we need to estimate the probability of detection for the whole taxon including uncertain identifications by fitting a detection function to the distance data. To do this we use detection functions which are as close as possible in terms of both structure and underlying data to the original models used to estimate the probabilities of detection for positively identified sightings. If the number of sightings for the whole taxon including uncertain identifications exceeds the appropriate threshold (30 or more sightings for this study), then the detection function used is derived from the common species detection function, otherwise it is derived from the rare species detection function. The detection functions used to estimate the probability of detection for the whole taxon include exactly the same covariates as the detection functions from which they were derived (e.g. in this study, for models derived from the common species detection function: species, cluster size, survey, observer and sea state; for models derived from the rare species detection function: body length<sup>2</sup>, behaviour, cluster size, survey, observer and

sea state. The underlying data includes all the data upon which the original detection function models are based (i.e. all sightings positively identified to species, across all species, including species outside the target taxon) plus those sightings from within the target taxon that were not identified to species. Sightings not identified to species from taxa outside the target taxon are not included.

For a target taxon with 30 or more sightings, where a model derived from the common species detection function is used to estimate the probability of detection, species is included explicitly as a covariate in the model. In this case, for all records within the underlying data belonging to the target taxon, including uncertain identifications, the species codes are changed from their original values to a single unique identifier for the taxon.

For a target taxon with less than 30 sightings, where a model derived from the rare species detection function is used to estimate the probability of detection, species is not explicitly included as a covariate within the model, but underlying covariates (e.g. Behaviour and body length) are used to capture the variation in the probability of detection between species. In this case, for positively identified sightings within the target taxon, the values of these covariates are left unchanged. For sightings within the target taxon not identified to species, the values of the quantitative covariates capturing variation between species (e.g. body length) are estimated separately for each survey as the mean across individual birds of that covariate based on positively identified sightings within the taxon. For qualitative covariates (e.g. Behaviour) usually a single value will apply across the whole taxon and the same value can be used for unidentified sightings as those identified to species. Should a situation arise where a qualitative variable varies between species within a taxon, then the value of the covariate for which most individuals were recorded could be used.

Following these changes to the data (i.e. adding unidentified sightings within the target taxon, and modifying the values of the covariates capturing the variation between species for the sightings within the taxon) the detection function is refitted. Along with estimates of the probability of detection for sightings positively identified to species outside the taxon, the resulting model provides estimates for the target taxon, including sightings not identified to species. This detection function provides the necessary probability of detection estimates required for the dht function to compute estimates of  $D_u$ , the density of birds in the target taxon, including birds not identified to species, along with the estimates of the standard error  $S_u$  and degrees of freedom  $V_u$  associated with this estimate. This is achieved by providing the dht function with this ddf model to estimate the probability of detection plus the appropriate subset of data as the set of observations upon which to basis its encounter rate estimates (i.e. all sightings on the water and in transect within the target taxon including sightings not identified to species)

#### **Conditions defining when corrections for uncertain identifications need to be applied**

Obviously, corrections only need to be applied if the particular area/season for which we are estimating density has any sightings within the taxon that were not identified to species, and furthermore the inclusion of these sightings increases the density estimate over that obtained if only sightings positively identified to species are considered (i.e.  $D_u > D_p$ ).

Where additional unidentified sightings do exist, it is not a foregone conclusion that their inclusion in the distance sampling analyses will increase the density estimate for the taxonomic group concerned. The reason for this is that from the perspective of fitting a detection function for a particular species failing to detect birds is indistinguishable from failing to identify them. Indeed these two phenomena are likely to behave similarly to one another, both increasing in likelihood with increase distance from the observer. Therefore, the fitting of a detection function for an individual species will potentially capture both failure to detect and failure to identify for birds belonging to that species. If the detection function

does capture both phenomena fully the inclusion of sightings not included to species will not increase the density estimate for the taxonomic group.

More formally, we need only correct density estimates for area/season combinations where both of the following criteria are satisfied:

$$n_U > n_P \quad (7)$$

and

$$\frac{n_U}{n_P} > \frac{P_U}{P_P} \quad (8)$$

Where:

- $n_U$  is the number of animals seen within the area/season including all sightings within the taxon.
- $n_P$  is the number of animals seen within the area/season only including those animals within the taxon that were positively identified to species.
- $P_U$  is the probability of detection for all sightings within the taxon, including those not identified to species.
- $P_P$  is the probability of detection for sightings identified to species within the taxon.

### **Correcting for uncertain identifications in hierarchically nested taxa**

So far we have considered only a single taxonomic group with  $n$  species and one category of unidentified individuals. However, the taxonomic groups within which we wish to apply corrections for unidentified sightings are often nested hierarchically, with a need to assign uncertain observations at each step in the hierarchy. For example within auks:

- There is a guillemot/razorbill taxonomic grouping, within which birds will have been positively identified as one of guillemot or razorbill, or identified as belonging to one or other of these two species, although it is uncertain which one.
- There is a puffin/little auk taxonomic grouping, within which birds will have been positively identified as one of puffin or little auk, or identified as belonging to one or other of these two species, although it is uncertain which one.
- There is an overall auk grouping including all the previous groupings (i.e. the individual species, plus the guillemot/razorbill grouping and the puffin/little auk grouping) and an additional “uncertain” group where the observer can be no more precise than saying the bird was an auk.

We handle such hierarchies by correcting for uncertainties in the smaller taxonomic groupings, where certainty is greatest, first, and then working upwards.

For example, for auks:

- We calculate corrected densities for guillemot and razorbill within the guillemot/razorbill grouping and corrected densities for puffin and little auk within the puffin/little auk grouping, as described above.
- Then, to apply the correction within the overall auk grouping we take as our initial uncorrected densities for individual species, and their associated variances, these estimates already corrected for uncertainty at lower (i.e. more precise) taxonomic levels. Thus the density of positively identified birds at the auk level is taken to include birds already assigned to lower taxonomic groupings, even if not to particular species.

### **Comparison with approach recommended by Maclean *et al.* (2009)**

To correct for uncertain identifications Maclean *et al.* (2009) recommend the following approach “the relative abundance of each of the species comprising the taxon is calculated from positively identified individuals. Individuals of the generic taxon can then be randomly assigned a species identity using the ratio of relative abundances to determine the total

number assigned to each species.” The approach adopted here is similar to this but differs in that whereas we assign uncertain observations to species on the basis of their densities derived from distance sampling analyses the Maclean *et al.* approach does it before any distance analysis is performed, on the basis of their sample sizes. Advantages of our approach over the Maclean *et al.* approach are:

- Our approach takes into account any differences between species in their probability of detection when estimating their relative abundance when assigning unidentified sightings to species. As the Maclean *et al.* approach is based upon raw sample sizes, it ignores any differences between species in this regard.
- Detection function modelling can potentially capture both failing to detect birds and failing to identify them to species. Our approach takes into account this potential for detection function modelling to capture the failure to identify sightings to species, whereas the Maclean *et al.* approach does not. This results in corrections either being applied less often, or being of a smaller magnitude than would result from the Maclean *et al.* approach.
- Our approach explicitly considers the effects of any correction on the variance of the corrected density, whereas the Maclean *et al.* approach does not.

The correction described in Equation 1 whereby the density estimate for each species is inflated by the same correction factor  $D_u/D_p$  corrects for any failure to identify species that has not already been captured by the detection functions based upon positively identified sightings. An example of where such detection function modelling might not fully capture any failure to identify birds would be if all birds close to the observer were detected, but only some could be assigned to species. Another example would be where if birds are more likely to be recorded as uncertain if (at a given distance from the observer) they occur in larger groups than sightings identified to species, or in mixed groups with a species composition differing to that found in groups identified to species. The correction applied inflates the number of birds assigned to each species by the same factor so as to leave no birds within the taxon unaccounted for, and effectively assumes all birds assigned to a species, are subject to the same probabilities of detection and identification as the positively identified individuals within that species.

For this project we obtain the estimate for the density of positively identified birds within the taxon  $D_p$ , by summing density estimates for individual species within that taxon. This approach allows density estimates to be combined from different distance sampling analyses, with the common species detection function being used for some analyses and the rare species detection function for others. Where taxa are hierarchically nested, it also allows density estimates for species corrected for uncertainty at a lower taxonomic level to be combined with density estimates for other species at a higher taxonomic level. However, where all the density estimates within a taxon originate from a single detection function model, and non are for species corrected for uncertain identification at a lower taxonomic level, then a better approach might be to directly estimate the density of positively identified sightings within the target taxon in a single call to the dht function with the appropriate subset of the observation data as the basis of its calculations (i.e. all positively identified sightings within the taxon). This avoids to having to make the simplifying assumption when estimating the standard error  $S_p$  and degrees of freedom  $V_p$  associated with the  $D_p$  estimate that the density estimates for the individual species are independent of one another, yielding more robust estimates. This more robust approach to estimating  $D_p$  and its associate standard error and degrees of freedom estimates could be adopted when ever all the density estimates originate from a single detection function model and either the taxon is not part of a hierarchy, or it is the lowest, most precise, taxon within this hierarchy.



Table 1: Survey effort achieved on each transect expressed as a percentage of transect length

Table 1a: For each survey date

Date	Transect											Overall
	2	4	6	8	10	12	14	16	18	20	22	
27/03/2012	103%	104%	99%	99%	96%	99%	101%	98%	100%	98%	97%	99%
18/04/2012	89%	103%	101%	100%	103%	94%	100%	97%	102%	99%	97%	99%
27/05/2012	107%	101%	102%	103%	104%	100%	100%	100%	101%	92%	102%	101%
30/06/2012	99%	95%	100%	95%	99%	101%	96%	100%	96%	97%	102%	98%
20/07/2012	95%	97%	95%	96%	101%	100%	100%	101%	99%	97%	103%	99%
20/08/2012	94%	93%	98%	93%	98%	87%	86%	98%	99%	100%		89%
11/12/2012	101%	102%	101%	100%	102%	97%	90%					56%
17/02/2013	105%	101%	103%	102%	91%	99%	101%	101%	102%	98%	91%	99%
04/03/2013	98%	100%	98%	97%	103%	94%	97%					56%
05/03/2013	107%	95%	101%	99%	98%	96%	96%	99%	100%	97%	92%	98%
30/03/2013	97%	105%	101%	100%	102%	89%	101%	100%	98%	97%	91%	98%
16/05/2013							71%					7%
03/06/2013	102%	104%	103%	101%	96%	96%	99%	98%	99%	100%	102%	100%
25/06/2013	111%	101%	102%	100%	97%	100%	99%	103%	100%	99%	100%	101%
10/07/2013	96%	95%	102%	100%	99%	98%	102%	98%	98%	101%	94%	99%
09/09/2013	106%	83%	99%	90%	98%	97%	102%	101%	97%	95%	94%	96%
23/10/2013	99%	89%	101%	102%	97%	101%	98%	62%	25%			68%
17/02/2014	76%	102%	93%	97%	88%							34%
19/02/2014									47%	98%	94%	22%
12/03/2014	94%	92%	92%	87%	93%	96%	95%					53%
Transect length (km)	2.7	5.1	6.4	7.7	7.4	8.4	7.5	11.0	10.0	8.3	5.0	79.5

Table 1b: For each survey

Survey	Transect											Overall
	2	4	6	8	10	12	14	16	18	20	22	
27/03/2012	103%	104%	99%	99%	96%	99%	101%	98%	100%	98%	97%	99%
18/04/2012	89%	103%	101%	100%	103%	94%	100%	97%	102%	99%	97%	99%
27/05/2012	107%	101%	102%	103%	104%	100%	100%	100%	101%	92%	102%	101%
30/06/2012	99%	95%	100%	95%	99%	101%	96%	100%	96%	97%	102%	98%
20/07/2012	95%	97%	95%	96%	101%	100%	100%	101%	99%	97%	103%	99%
20/08/2012	94%	93%	98%	93%	98%	87%	86%	98%	99%	100%		89%
11/12/2012	101%	102%	101%	100%	102%	97%	90%					56%
17/02/2013	105%	101%	103%	102%	91%	99%	101%	101%	102%	98%	91%	99%
04/03/2013	98%	100%	98%	97%	103%	94%	97%					56%
05/03/2013	107%	95%	101%	99%	98%	96%	96%	99%	100%	97%	92%	98%
30/03/2013	97%	105%	101%	100%	102%	89%	101%	100%	98%	97%	91%	98%
03/06/2013	102%	104%	103%	101%	96%	96%	99%	98%	99%	100%	102%	100%
25/06/2013	111%	101%	102%	100%	97%	100%	99%	103%	100%	99%	100%	101%
10/07/2013	96%	95%	102%	100%	99%	98%	102%	98%	98%	101%	94%	99%
09/09/2013	106%	83%	99%	90%	98%	97%	102%	101%	97%	95%	94%	96%
23/10/2013	99%	89%	101%	102%	97%	101%	98%	62%	25%			68%
18/02/2014	76%	102%	93%	97%	88%				47%	98%	94%	56%
12/03/2014	94%	92%	92%	87%	93%	96%	95%					53%
Transect length (km)	2.7	5.1	6.4	7.7	7.4	8.4	7.5	11.0	10.0	8.3	5.0	79.5

Table 2: Number of observations (clusters) of each species on each survey

Table 2a: Birds on the water

Survey	Red-throated diver	Black-throated diver	Great northern diver	Fulmar	Mann shearwater	Storm petrel	Gannet	Cormorant	Shag	Arctic skua	Great skua	Common gull	Herring gull	Great black-backed gull	Kitiwake	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	auk sp.	Guillemot/ razorbill	Total Without uncertain ID	Total with uncertain ID
27/03/2012	1	0	0	105	0	0	0	0	8	0	1	0	0	1	0	0	37	3	0	0	41	0	0	197	197
18/04/2012	0	0	0	41	0	0	2	0	3	0	1	0	0	0	3	0	16	23	1	0	14	0	1	104	105
27/05/2012	0	0	0	196	0	0	11	0	4	2	15	0	0	0	7	1	374	49	2	0	102	2	23	763	788
30/06/2012	0	0	0	16	1	0	2	0	0	0	2	0	0	0	6	0	35	17	0	0	51	1	4	130	135
20/07/2012	0	0	0	30	0	0	1	0	8	1	1	0	0	0	0	0	14	8	0	0	97	0	0	160	160
20/08/2012	0	0	0	28	0	1	12	0	0	0	2	0	0	0	0	0	0	7	0	0	4	0	0	54	54
11/12/2012	0	0	0	1	0	0	1	0	3	0	0	1	1	1	0	0	1	0	0	0	0	0	0	9	9
17/02/2013	0	0	0	27	0	0	0	0	6	0	0	0	1	2	2	0	26	1	2	0	0	0	0	67	67
04/03/2013	0	0	0	12	0	0	0	0	9	0	0	0	0	2	1	0	36	0	0	0	0	0	0	60	60
05/03/2013	0	0	0	31	0	0	0	0	10	0	0	0	0	1	0	0	91	4	1	1	4	0	0	143	143
30/03/2013	0	0	0	25	0	0	0	0	5	0	0	0	0	1	0	0	18	6	1	0	14	1	1	70	72
03/06/2013	0	0	0	24	0	0	1	0	1	0	1	0	0	0	5	0	25	33	1	0	58	0	0	149	149
25/06/2013	0	0	1	33	0	0	0	0	2	0	2	1	0	0	5	0	46	19	0	0	64	0	2	173	175
10/07/2013	0	0	0	91	0	0	0	0	6	0	3	0	0	0	1	0	153	16	2	0	40	0	1	312	313
09/09/2013	0	0	0	37	0	0	12	0	2	0	4	0	0	0	0	0	3	1	1	0	4	0	0	64	64
23/10/2013	0	0	0	0	0	0	4	1	2	0	0	0	0	0	3	0	8	3	0	0	2	0	0	23	23
18/02/2014	0	0	0	19	0	0	0	0	2	0	0	0	0	0	0	0	3	0	1	1	0	0	0	26	26
12/03/2014	0	1	0	13	0	0	0	0	2	0	0	0	2	3	0	0	9	1	0	0	2	0	0	33	33
Total	1	1	1	729	1	1	46	1	73	3	32	2	4	11	33	1	895	191	12	2	497	4	32	2537	2573

Table 2b: Birds in flight

Survey	Red-throated diver	Fulmar	Sooty shearwater	Manx shearwater	Storm petrel	Gannet	Shag	Arctic skua	Great skua	Black-headed gull	Common gull	Herring gull	Great black-backed gull	Kitiwake	Common tern	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	large gull sp. (HG, LB or GB)	Guillemot/razorb ill	Total without uncertain ID	Total with uncertain ID
27/03/2012	0	28	0	0	0	3	0	0	0	0	0	1	0	2	0	0	7	3	0	0	7	0	2	51	53
18/04/2012	0	49	0	0	0	7	0	1	1	0	1	0	0	4	0	0	14	8	0	0	1	0	2	86	88
27/05/2012	0	76	0	0	0	6	0	0	9	0	0	0	1	13	0	2	66	14	0	0	23	0	5	210	215
30/06/2012	0	65	0	3	3	9	0	0	12	0	0	0	1	41	0	2	22	17	0	0	15	0	1	190	191
20/07/2012	0	113	0	1	0	8	1	0	10	0	0	0	0	30	0	0	9	7	1	0	24	0	0	204	204
20/08/2012	0	80	0	0	1	8	3	1	18	0	0	0	1	1	0	0	0	0	0	0	0	0	0	113	113
11/12/2012	0	36	0	0	0	3	3	0	0	0	2	0	3	1	0	0	0	3	1	0	0	0	0	52	52
17/02/2013	0	49	0	0	0	1	2	0	0	0	0	0	1	8	0	0	16	2	0	0	0	1	1	79	81
04/03/2013	0	19	0	0	0	3	0	0	0	0	0	0	1	4	0	0	5	0	0	0	1	0	0	33	33
05/03/2013	0	44	0	0	0	5	0	0	0	0	1	1	0	11	0	0	11	2	0	0	1	0	2	76	78
30/03/2013	0	30	0	0	0	2	0	0	0	0	3	0	0	16	0	0	4	2	0	0	0	0	0	57	57
03/06/2013	0	49	0	0	1	2	1	1	8	0	0	0	0	7	0	3	19	11	0	0	9	0	0	111	111
25/06/2013	0	52	0	0	3	8	1	0	5	1	2	0	0	28	1	2	36	27	1	0	22	0	1	189	190
10/07/2013	0	55	0	1	4	7	1	2	6	0	1	0	0	20	0	1	37	16	0	0	10	0	0	161	161
09/09/2013	0	59	1	0	0	43	1	0	6	0	0	0	0	4	0	0	0	1	0	0	0	0	0	115	115
23/10/2013	0	6	0	0	0	19	0	0	0	0	0	1	3	11	0	0	0	1	0	0	1	0	0	42	42
18/02/2014	1	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	37	37
12/03/2014	1	28	0	0	0	0	0	0	0	0	0	0	0	2	0	0	8	0	0	0	0	0	1	39	40
Total	2	872	1	5	12	134	13	5	75	1	10	3	11	203	1	10	255	114	3	1	114	1	15	1845	1861

Table 3: Total number of individual animals of each species seen on each survey

Table 3a: Birds on water

Survey	Red-throated diver	Black-throated diver	Great northern diver	Fulmar	Manx shearwater	Storm petrel	Gannet	Cormorant	Shag	Arctic skua	Great skua	Common gull	Herring gull	Great black-backed gull	Kitiwake	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	auk sp.	Guillemot/razorbill	Total without uncertain ID	Total with uncertain ID
27/03/2012	1	0	0	690	0	0	0	0	11	0	1	0	0	1	0	0	60	5	0	0	48	0	0	817	817
18/04/2012	0	0	0	121	0	0	2	0	6	0	1	0	0	0	123	0	23	50	2	0	22	0	7	350	357
27/05/2012	0	0	0	1134	0	0	37	0	4	3	89	0	0	0	34	8	1774	92	3	0	250	3	818	3428	4249
30/06/2012	0	0	0	40	3	0	2	0	0	0	2	0	0	0	172	0	65	51	0	0	61	10	44	396	450
20/07/2012	0	0	0	43	0	0	1	0	13	1	1	0	0	0	0	0	27	21	0	0	188	0	0	295	295
20/08/2012	0	0	0	41	0	1	17	0	0	0	4	0	0	0	0	0	0	7	0	0	4	0	0	74	74
11/12/2012	0	0	0	20	0	0	1	0	4	0	0	1	1	1	0	0	1	0	0	0	0	0	0	29	29
17/02/2013	0	0	0	213	0	0	0	0	9	0	0	0	1	2	2	0	31	1	2	0	0	0	0	261	261
04/03/2013	0	0	0	41	0	0	0	0	17	0	0	0	0	2	1	0	51	0	0	0	0	0	0	112	112
05/03/2013	0	0	0	151	0	0	0	0	21	0	0	0	0	1	0	0	203	11	1	1	4	0	0	393	393
30/03/2013	0	0	0	162	0	0	0	0	6	0	0	0	0	1	0	0	19	6	1	0	15	1	2	210	213
03/06/2013	0	0	0	41	0	0	1	0	17	0	1	0	0	0	70	0	41	91	1	0	80	0	0	343	343
25/06/2013	0	0	1	56	0	0	0	0	15	0	2	2	0	0	755	0	118	306	0	0	84	0	3	1339	1342
10/07/2013	0	0	0	129	0	0	0	0	10	0	6	0	0	0	5	0	218	41	2	0	46	0	1	457	458
09/09/2013	0	0	0	90	0	0	16	0	2	0	5	0	0	0	0	0	5	1	1	0	4	0	0	124	124
23/10/2013	0	0	0	0	0	0	9	2	2	0	0	0	0	0	7	0	11	5	0	0	2	0	0	38	38
18/02/2014	0	0	0	66	0	0	0	0	2	0	0	0	0	0	0	0	6	0	1	1	0	0	0	76	76
12/03/2014	0	1	0	36	0	0	0	0	4	0	0	0	3	3	0	0	9	1	0	0	2	0	0	59	59
Total	1	1	1	3074	3	1	86	2	143	4	112	3	5	11	1169	8	2662	689	14	2	810	14	875	8801	9690

Table 3b: Birds in flight

Survey	Red-throated diver	Fulmar	Sooty shearwater	Manx shearwater	Storm petrel	Gannet	Shag	Arctic skua	Great skua	Black-headed gull	Common gull	Herring gull	Great black-backed gull	Kitiwake	Common tern	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	large gull sp. (HG, LB or GB)	Guillemot/razorbill	Total without uncertain ID	Total with uncertain ID
27/03/2012	0	33	0	0	0	3	0	0	0	0	0	1	0	2	0	0	15	14	0	0	13	0	3	81	84
18/04/2012	0	67	0	0	0	7	0	1	1	0	1	0	0	6	0	0	34	16	0	0	2	0	3	135	138
27/05/2012	0	92	0	0	0	7	0	0	9	0	0	0	1	20	0	26	169	23	0	0	83	0	17	430	447
30/06/2012	0	74	0	7	3	11	0	0	13	0	0	0	1	854	0	3	26	24	0	0	18	0	2	1034	1036
20/07/2012	0	136	0	1	0	8	1	0	10	0	0	0	0	179	0	0	13	16	1	0	42	0	0	407	407
20/08/2012	0	95	0	0	1	10	5	1	19	0	0	0	1	1	0	0	0	0	0	0	0	0	0	133	133
11/12/2012	0	49	0	0	0	5	3	0	0	0	2	0	4	1	0	0	0	6	1	0	0	0	0	71	71
17/02/2013	0	137	0	0	0	1	2	0	0	0	0	0	1	8	0	0	23	2	0	0	0	1	4	174	179
04/03/2013	0	19	0	0	0	4	0	0	0	0	0	0	1	4	0	0	10	0	0	0	1	0	0	39	39
05/03/2013	0	54	0	0	0	5	0	0	0	0	1	1	0	17	0	0	14	4	0	0	1	0	2	97	99
30/03/2013	0	34	0	0	0	4	0	0	0	0	3	0	0	27	0	0	7	5	0	0	0	0	0	80	80
03/06/2013	0	52	0	0	1	2	1	1	9	0	0	0	0	7	0	5	59	21	0	0	20	0	0	178	178
25/06/2013	0	66	0	0	3	9	1	0	7	1	6	0	0	128	2	6	62	39	1	0	45	0	1	376	377
10/07/2013	0	64	0	2	4	14	1	2	6	0	1	0	0	41	0	6	52	17	0	0	16	0	0	226	226
09/09/2013	0	82	1	0	0	74	1	0	6	0	0	0	0	4	0	0	0	1	0	0	0	0	0	169	169
23/10/2013	0	9	0	0	0	31	0	0	0	0	0	1	3	15	0	0	0	1	0	0	4	0	0	64	64
18/02/2014	1	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	59	59
12/03/2014	1	36	0	0	0	0	0	0	0	0	0	0	0	2	0	0	12	0	0	0	0	0	3	51	54
Total	2	1155	1	10	12	195	15	5	80	1	14	3	12	1316	2	46	497	189	3	1	245	1	35	3804	3840

Table 4: Mean cluster size of each species on each survey

Table 4a: Birds on water

Survey	Red-throated diver	Black-throated diver	Great northern diver	Fulmar	Manx shearwater	Storm petrel	Gannet	Cormorant	Shag	Arctic skua	Great skua	Common gull	Herring gull	Great black-backed gull	Kitiwake	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	auk sp.	Guillemot/razorbill
27/03/2012	1.0			6.6					1.4		1.0			1.0			1.6	1.7			1.2		
18/04/2012				3.0			1.0		2.0		1.0				41.0		1.4	2.2	2.0		1.6		7.0
27/05/2012				5.8			3.4		1.0	1.5	5.9				4.9	8.0	4.7	1.9	1.5		2.5	1.5	35.6
30/06/2012				2.5	3.0		1.0				1.0				28.7		1.9	3.0			1.2	10.0	11.0
20/07/2012				1.4			1.0		1.6	1.0	1.0						1.9	2.6			1.9		
20/08/2012				1.5		1.0	1.4				2.0							1.0			1.0		
11/12/2012				20.0			1.0		1.3			1.0	1.0	1.0			1.0						
17/02/2013				7.9					1.5				1.0	1.0	1.0		1.2	1.0	1.0				
04/03/2013				3.4					1.9					1.0	1.0		1.4						
05/03/2013				4.9					2.1					1.0			2.2	2.8	1.0	1.0	1.0		
30/03/2013				6.5					1.2					1.0			1.1	1.0	1.0		1.1	1.0	2.0
03/06/2013				1.7			1.0		17.0		1.0				14.0		1.6	2.8	1.0		1.4		
25/06/2013			1.0	1.7					7.5		1.0	2.0			151.0		2.6	16.1			1.3		1.5
10/07/2013				1.4					1.7		2.0				5.0		1.4	2.6	1.0		1.2		1.0
09/09/2013				2.4			1.3		1.0		1.3						1.7	1.0	1.0		1.0		
23/10/2013							2.3	2.0	1.0						2.3		1.4	1.7			1.0		
18/02/2014				3.5					1.0								2.0		1.0	1.0			
12/03/2014		1.0		2.8					2.0				1.5	1.0			1.0	1.0			1.0		
Total	1.0	1.0	1.0	4.2	3.0	1.0	1.9	2.0	2.0	1.3	3.5	1.5	1.3	1.0	35.4	8.0	3.0	3.6	1.2	1.0	1.6	3.5	27.3

Table 4b: Birds in flight

Survey	Red-throated diver	Fulmar	Sooty shearwater	Manx shearwater	Storm petrel	Gannet	Shag	Arctic skua	Great skua	Black-headed gull	Common gull	Herring gull	Great black-backed gull	Kitiwake	Common tern	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	large gull sp. (HG, LB or GB)	Guillemot/ razorbill
27/03/2012		1.2				1.0						1.0		1.0			2.1	4.7			1.9		1.5
18/04/2012		1.4				1.0		1.0	1.0		1.0			1.5			2.4	2.0			2.0		1.5
27/05/2012		1.2				1.2			1.0				1.0	1.5		13.0	2.6	1.6			3.6		3.4
30/06/2012		1.1		2.3	1.0	1.2			1.1				1.0	20.8		1.5	1.2	1.4			1.2		2.0
20/07/2012		1.2		1.0		1.0	1.0		1.0					6.0			1.4	2.3	1.0		1.8		
20/08/2012		1.2			1.0	1.3	1.7	1.0	1.1				1.0	1.0									
11/12/2012		1.4				1.7	1.0				1.0		1.3	1.0				2.0	1.0				
17/02/2013		2.8				1.0	1.0						1.0	1.0			1.4	1.0				1.0	4.0
04/03/2013		1.0				1.3							1.0	1.0			2.0				1.0		
05/03/2013		1.2				1.0					1.0	1.0		1.5			1.3	2.0			1.0		1.0
30/03/2013		1.1				2.0					1.0			1.7			1.8	2.5					
03/06/2013		1.1			1.0	1.0	1.0	1.0	1.1					1.0		1.7	3.1	1.9			2.2		
25/06/2013		1.3			1.0	1.1	1.0		1.4	1.0	3.0			4.6	2.0	3.0	1.7	1.4	1.0		2.0		1.0
10/07/2013		1.2		2.0	1.0	2.0	1.0	1.0	1.0		1.0			2.1		6.0	1.4	1.1			1.6		
09/09/2013		1.4	1.0			1.7	1.0		1.0					1.0				1.0					
23/10/2013		1.5				1.6						1.0	1.0	1.4				1.0			4.0		
18/02/2014	1.0	1.6															1.0			1.0			
12/03/2014	1.0	1.3												1.0			1.5						3.0
Total	1.0	1.3	1.0	2.0	1.0	1.5	1.2	1.0	1.1	1.0	1.4	1.0	1.1	6.5	2.0	4.6	1.9	1.7	1.0	1.0	2.1	1.0	2.3



Table 5: Median cluster size of each species on each survey

Table 5a: Birds on water

Survey	Red-throated diver	Black-throated diver	Great northern diver	Fulmar	Mannx shearwater	Storm petrel	Gannet	Cormorant	Shag	Arctic skua	Great skua	Common gull	Herring gull	Great black-backed gull	Kitiwake	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	auk sp.	Guillemot/ razorbill		
27/03/2012	1.0			3.0					1.0		1.0			1.0			1.0	1.0			1.0				
18/04/2012				1.0			1.0		2.0		1.0				40.0		1.0	2.0	2.0		1.0			7.0	
27/05/2012				2.0			2.0		1.0	1.5	7.0				5.0	8.0	2.0	1.0	1.5		2.0	1.5		30.0	
30/06/2012				1.0	3.0		1.0				1.0				35.0		1.0	1.0			1.0	10.0		11.5	
20/07/2012				1.0			1.0		1.5	1.0	1.0						1.0	2.0			1.0				
20/08/2012				1.0		1.0	1.0				2.0							1.0			1.0				
11/12/2012				20.0			1.0		1.0			1.0	1.0	1.0			1.0								
17/02/2013				1.0					1.0				1.0	1.0	1.0		1.0	1.0	1.0						
04/03/2013				1.0					2.0					1.0	1.0		1.0								
05/03/2013				1.0					2.0					1.0			1.0	3.0	1.0	1.0	1.0				
30/03/2013				1.0					1.0					1.0			1.0	1.0	1.0		1.0	1.0		2.0	
03/06/2013				1.0			1.0		17.0		1.0				12.0		1.0	2.0	1.0		1.0				
25/06/2013			1.0	1.0					7.5		1.0	2.0			150.0		1.0	3.0			1.0			1.5	
10/07/2013				1.0					1.0		2.0				5.0		1.0	1.0	1.0		1.0			1.0	
09/09/2013				1.0			1.0		1.0		1.0						1.0	1.0	1.0		1.0				
23/10/2013							1.0	2.0	1.0						1.0		1.0	2.0			1.0				
18/02/2014				1.0					1.0								2.0		1.0	1.0					
12/03/2014		1.0		1.0					2.0				1.5	1.0			1.0	1.0			1.0				
Total	1.0	1.0	1.0	1.0	3.0	1.0	1.0	2.0	1.0	1.0	2.0	1.5	1.0	1.0	8.0	8.0	1.0	1.0	1.0	1.0	1.0	1.5		20.0	

Table 5b: Birds in flight

Survey	Red-throated diver	Fulmar	Sooty shearwater	Mannx shearwater	Storm petrel	Gannet	Shag	Arctic skua	Great skua	Black-headed gull	Common gull	Herring gull	Great black-backed gull	Kitiwake	Common tern	Arctic tern	Guillemot	Razorbill	Black guillemot	Little Auk	Puffin	large gull sp. (HG, LB or GB)	Guillemot/razorbill
27/03/2012		1.0				1.0						1.0		1.0			1.0	2.0			2.0		1.5
18/04/2012		1.0				1.0		1.0	1.0		1.0			1.5			2.0	2.0			2.0		1.5
27/05/2012		1.0				1.0			1.0				1.0	1.0		13.0	1.0	1.0			3.0		3.0
30/06/2012		1.0		3.0	1.0	1.0			1.0				1.0	1.0		1.5	1.0	1.0			1.0		2.0
20/07/2012		1.0		1.0		1.0	1.0		1.0					1.0			1.0	1.0	1.0		1.0		
20/08/2012		1.0			1.0	1.0	1.0	1.0	1.0				1.0	1.0									
11/12/2012		1.0				1.0	1.0				1.0		1.0	1.0				1.0	1.0				
17/02/2013		1.0				1.0	1.0						1.0	1.0			1.0	1.0				1.0	4.0
04/03/2013		1.0				1.0							1.0	1.0			2.0				1.0		
05/03/2013		1.0				1.0					1.0	1.0		1.0			1.0	2.0			1.0		1.0
30/03/2013		1.0				2.0					1.0			1.0			1.0	2.5					
03/06/2013		1.0			1.0	1.0	1.0	1.0	1.0					1.0		2.0	2.0	1.0			2.0		
25/06/2013		1.0			1.0	1.0	1.0		1.0	1.0	3.0			1.0	2.0	3.0	1.0	1.0	1.0		1.5		1.0
10/07/2013		1.0		2.0	1.0	1.0	1.0	1.0	1.0		1.0			1.0		6.0	1.0	1.0			1.5		
09/09/2013		1.0	1.0			1.0	1.0		1.0					1.0				1.0					
23/10/2013		1.0				1.0						1.0	1.0	1.0				1.0			4.0		
18/02/2014	1.0	1.0															1.0			1.0			
12/03/2014	1.0	1.0												1.0			1.0						3.0
Total	1.0	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0

Table 6: Goodness of fit of base model for predicting probability of detection for common species.

ID	Model	AIC Term 1	AIC Term 2	AIC	$\Delta$ AIC Term 1	$\Delta$ AIC Term 2	$\Delta$ AIC	Comment
1	1	6901.9	2	6903.9	0.0	0	0.0	Null model
2	log(size)	6853.5	4	6857.5	-48.4	2	-46.4	
3	survey	6788.0	36	6824.0	-113.9	34	-79.9	
4	species	6778.5	18	6796.5	-123.4	16	-107.4	
5	log(size) + survey	6751.7	38	6789.7	-150.3	36	-114.3	
6	species + log(size)	6730.3	20	6750.3	-171.7	18	-153.7	
7	species + survey	6681.0	52	6733.0	-220.9	50	-170.9	
8	species + log(size) + survey	6649.3	54	6703.3	-252.7	52	-200.7	Base model

Table 7: Goodness of fit of alternative detection function models for predicting probability of detection for common species.

ID	Model	AIC Term 1	AIC Term 2	AIC	$\Delta$ AIC Term 1	$\Delta$ AIC Term 2	$\Delta$ AIC	Comment
1	1	6901.9	2	6903.9	0.0	0	0.0	Null model
2	species + log(size) + windforce	6722.3	22	6744.3	-179.7	20	-159.7	
3	species + log(size) + swellheight	6716.5	22	6738.5	-185.4	20	-165.4	
4	species + log(size) + observer	6684.6	38	6722.6	-217.3	36	-181.3	
5	species + log(size) + observer + swellheight	6677.4	40	6717.4	-224.5	38	-186.5	
6	species + log(size) + observer + windforce	6673.5	40	6713.5	-228.4	38	-190.4	
7	species + log(size) + seastate	6690.1	22	6712.1	-211.8	20	-191.8	
8	species + log(size) + seastate + swellheight	6685.1	24	6709.1	-216.8	22	-194.8	
9	species + log(size) + survey	6649.3	54	6703.3	-252.7	52	-200.7	Base model
10	species + log(size) + survey + windforce	6643.5	56	6699.5	-258.4	54	-204.4	
11	species + log(size) + survey + swellheight	6641.6	56	6697.6	-260.4	54	-206.4	
12	species + log(size) + survey + observer	6623.0	72	6695.0	-279.0	70	-209.0	
13	species + log(size) + survey + observer + windforce	6620.6	74	6694.6	-281.3	72	-209.3	
14	species + log(size) + survey + observer + swellheight	6618.0	74	6692.0	-283.9	72	-211.9	
15	species + log(size) + survey + seastate	6626.8	56	6682.8	-275.2	54	-221.2	
16	species + log(size) + survey + observer + seastate	6607.1	74	6681.1	-294.9	72	-222.9	
17	species + log(size) + observer + seastate + swellheight	6635.0	42	6677.0	-267.0	40	-227.0	
18	species + log(size) + observer + seastate	6635.0	40	6675.0	-266.9	38	-228.9	

Table 8: Coefficients for common species detection function model

Coefficient	Estimate	SE	CV
Intercept (Fulmar,observer 1)	5.84	0.12	2%
species Gannet	0.95	0.97	102%
species Great skua	0.39	0.66	168%
species Guillemot	-0.45	0.08	17%
species Kittiwake	-0.68	0.19	28%
species Other	-0.11	0.22	199%
species Puffin	-0.52	0.08	16%
species Razorbill	-0.68	0.09	14%
species Shag	-0.36	0.13	37%
log(size)	0.20	0.04	20%
observer 2	-0.42	0.09	22%
observer 3	-0.40	0.13	33%
observer 4	-0.27	0.09	32%
observer 5	-0.07	0.34	511%
observer 6	0.01	0.30	2613%
observer 7	-0.09	0.09	102%
observer 8	-0.43	0.08	18%
observer 9	-0.19	0.26	135%
observer 10	-0.52	0.51	98%
seastate	-0.17	0.02	14%

Table 9: Average and standardized (i.e. single bird, observer 8, sea state 2) probability of detection for each species from both the common species and rare species detection function models

species	n	JNCC <sup>1</sup>	Common species detection function			Rare species detection function				
			Average probability of detection		Standardised probability of detection for single individual	Body <sup>2</sup> length (m)	Behaviour	Average probability of detection		Standardised probability of detection for single individual
			Clusters	Individuals				Clusters	Individuals	
Red-throated diver	1	76.9%				0.61	Dive	59.4%	59.4%	46.3%
Black-throated diver	1	76.9%				0.66	Dive	59.4%	59.4%	48.2%
Great northern diver	1	76.9%				0.80	Dive	62.4%	62.4%	54.2%
Fulmar	729	90.9%	71.9%	82.1%	62.4%	0.48	Surface/Aerial	71.9%	81.9%	62.4%
Manx shearwater	1	76.9%				0.34	Surface/Aerial	52.7%	52.7%	58.9%
Storm petrel	1	66.7%				0.16	Surface/Aerial	57.0%	57.0%	56.2%
Gannet	46	100.0%	92.7%	94.1%	91.8%	0.94	Surface/Aerial	82.0%	85.2%	79.5%
Cormorant	1	90.9%				0.90	Dive	57.6%	57.6%	59.4%
Shag	73	90.9%	54.5%	58.9%	46.0%	0.72	Dive	59.0%	63.3%	50.6%
Arctic skua	3	76.9%				0.44	Surface/Aerial	77.1%	76.5%	61.3%
Great skua	32	76.9%	85.4%	89.9%	78.4%	0.56	Surface/Aerial	75.0%	82.0%	64.8%
Common gull	2	71.4%				0.41	Surface/Aerial	67.0%	69.2%	60.5%
Herring gull	4	71.4%				0.60	Surface/Aerial	68.0%	70.2%	66.2%
Great black-backed gull	11	71.4%				0.71	Surface/Aerial	74.2%	74.2%	70.2%
Kittiwake	33	71.4%	49.3%	72.9%	33.6%	0.39	Surface/Aerial	75.3%	90.0%	60.0%
Arctic tern	1	58.8%				0.34	Surface/Aerial	82.5%	82.5%	58.9%
Guillemot	895	71.4%	52.5%	63.1%	42.2%	0.40	Dive	51.0%	61.6%	40.6%
Razorbill	191	66.7%	41.3%	53.3%	33.6%	0.38	Dive	48.8%	60.9%	40.2%
Black guillemot	12	55.6%				0.31	Dive	42.6%	43.0%	38.9%
Little Auk	2	52.6%				0.18	Dive	43.6%	43.6%	37.4%
Puffin	497	66.7%	45.8%	50.1%	39.3%	0.28	Dive	45.1%	49.5%	38.5%
Other	41		63.0%	64.7%	57.2%					

<sup>1</sup> JNCC correction factors all taken from Stone et al. (1995), apart from black guillemot, taken from Skov et al. (1995)

<sup>2</sup> Body lengths taken from <http://www.bto.org/about-birds/birdfacts>

Table 10: Average probability of detection for different cluster size categories from the common species detection function model.

Cluster size category	N	Average probability of detection	
		Clusters	Individuals
1	1529	50.4%	50.4%
>1	1008	63.6%	72.8%
>5	294	75.7%	78.6%
>10	122	78.0%	79.9%
>20	53	80.3%	81.3%
>30	30	82.0%	82.2%

Table 11: Standardised (i.e. guillemot or kittiwake, observer 8, sea state 2) probability of detection for different cluster size categories from the common species detection function model.

Cluster size	Probability of detection	
	Guillemot	Kittiwake
1	42.2%	33.6%
2	48.2%	38.7%
3	51.9%	41.9%
4	54.6%	44.3%
5	56.7%	46.2%
10	63.2%	52.5%
20	69.3%	59.0%
30	72.7%	62.8%
40	74.9%	65.4%
50	76.6%	67.4%
100	81.3%	73.2%
200	85.2%	78.3%
300	87.1%	81.0%
400	88.4%	82.7%

Table 12: Average and standardised (i.e. single guillemot, observer 8) probability of detection for each sea state from the common species detection function model.

Sea state	N	Average probability of detection		Standardised probability of detection for a single guillemot
		Clusters	Individuals	
0	37	79.0%	90.5%	57.4%
1	582	59.7%	68.0%	49.6%
2	995	56.8%	69.0%	42.2%
3	644	55.5%	70.7%	35.7%
4	256	41.5%	57.0%	30.1%
5	23	34.6%	52.4%	25.4%

Table 13: Average and standardised (i.e. single guillemot, sea state 2) probability of detection for each observer from the common species detection function model.

Observer	N	Average probability of detection		Standardised probability of detection for a single guillemot
		Clusters	Individuals	
observer 1	489	66.2%	76.2%	61.8%
observer 2	346	52.3%	61.3%	42.8%
observer 3	46	49.7%	54.3%	43.6%
observer 4	460	54.6%	61.0%	49.3%
observer 5	32	70.5%	76.1%	58.7%
observer 6	25	68.6%	78.2%	62.3%
observer 7	498	62.0%	72.8%	57.8%
observer 8	606	46.1%	64.2%	42.2%
observer 9	26	41.8%	47.0%	52.8%
observer 10	9	40.8%	50.8%	38.8%

Table 14: Average probability of detection for each survey from the common species detection function model.

Survey	N	Average probability of detection	
		Clusters	Individuals
27/03/2012	197	59.6%	77.3%
18/04/2012	104	49.8%	62.7%
27/05/2012	763	62.9%	73.7%
30/06/2012	130	35.7%	46.4%
20/07/2012	160	54.7%	56.9%
20/08/2012	54	57.1%	62.1%
11/12/2012	9	54.9%	69.9%
17/02/2013	67	49.6%	70.7%
04/03/2013	60	48.6%	55.2%
05/03/2013	143	49.2%	59.7%
30/03/2013	70	53.2%	71.2%
03/06/2013	149	50.4%	55.5%
25/06/2013	173	50.3%	71.2%
10/07/2013	312	58.5%	60.5%
09/09/2013	64	68.7%	75.3%
23/10/2013	23	38.4%	42.5%
18/02/2014	26	59.7%	70.0%
12/03/2014	33	69.3%	76.2%



Table 15: Average probability of detection of clusters of each species on each survey from the common species detection function model.

Survey	Species								
	Fulmar	Gannet	Shag	Great skua	Kittiwake	Guillemot	Razorbill	Puffin	Other
27/03/2012	74.5%		60.8%	78.4%		53.6%	36.9%	43.2%	65.1%
18/04/2012	61.3%	95.7%	40.2%	86.0%	70.5%	46.1%	40.6%	41.3%	55.8%
27/05/2012	78.3%	95.3%	54.0%	89.9%	52.7%	60.5%	46.3%	55.4%	73.6%
30/06/2012	52.9%	86.7%		64.7%	53.1%	34.6%	30.1%	32.3%	51.7%
20/07/2012	75.8%	96.4%	73.3%	89.7%		59.0%	48.9%	48.9%	75.9%
20/08/2012	62.6%	91.8%		80.4%			31.2%	39.7%	57.8%
11/12/2012	84.8%	91.8%	42.9%			42.8%			62.4%
17/02/2013	68.5%		48.1%		33.8%	40.4%	34.1%		50.1%
04/03/2013	66.9%		50.2%		33.4%	44.4%			56.9%
05/03/2013	67.6%		50.2%			45.4%	39.7%	41.2%	60.6%
30/03/2013	74.2%		58.0%			46.7%	38.1%	43.9%	65.4%
03/06/2013	66.3%	94.2%	83.2%	84.0%	54.9%	51.6%	42.9%	48.3%	72.0%
25/06/2013	69.0%		63.8%	80.9%	72.7%	48.1%	46.5%	44.4%	67.9%
10/07/2013	74.3%		64.0%	86.3%	53.7%	55.5%	45.5%	48.6%	65.6%
09/09/2013	70.6%	94.5%	53.4%	88.3%		45.6%	34.1%	46.1%	65.6%
23/10/2013		84.4%	46.0%		26.6%	35.7%	30.5%	36.1%	55.8%
18/02/2014	67.2%		45.6%			40.0%			59.0%
12/03/2014	81.2%		67.2%			59.7%	51.1%	56.8%	74.1%

Table 16: Average probability of detection of individuals of each species on each survey from the common species detection function model.

Survey	Species								
	Fulmar	Gannet	Shag	Great skua	Kittiwake	Guillemot	Razorbill	Puffin	Other
27/03/2012	84.9%		61.6%	78.4%		57.2%	43.9%	45.4%	65.1%
18/04/2012	70.3%	95.7%	43.9%	86.0%	83.2%	47.5%	41.9%	43.9%	55.8%
27/05/2012	86.3%	96.5%	54.0%	92.0%	55.9%	70.3%	49.0%	61.6%	77.2%
30/06/2012	59.7%	86.7%		64.7%	58.1%	40.0%	38.6%	33.1%	51.7%
20/07/2012	76.8%	96.4%	75.5%	89.7%		63.9%	53.6%	52.2%	75.9%
20/08/2012	66.5%	92.6%		80.4%			31.2%	39.7%	57.8%
11/12/2012	84.8%	91.8%	41.5%			42.8%			62.4%
17/02/2013	83.0%		50.1%		33.8%	41.0%	34.1%		50.1%
04/03/2013	77.0%		52.6%		33.4%	46.0%			56.9%
05/03/2013	80.5%		53.2%			52.0%	42.0%	41.2%	60.6%
30/03/2013	84.9%		57.0%			46.8%	38.1%	43.9%	65.4%
03/06/2013	68.5%	94.2%	83.2%	84.0%	61.5%	56.1%	48.9%	49.7%	72.0%
25/06/2013	72.1%		64.9%	80.9%	79.9%	59.9%	68.4%	45.8%	68.9%
10/07/2013	76.0%		66.9%	86.5%	53.7%	58.0%	50.0%	49.1%	65.6%
09/09/2013	78.3%	94.7%	53.4%	88.8%		48.1%	34.1%	46.1%	65.6%
23/10/2013		87.8%	46.0%		32.6%	38.4%	32.3%	36.1%	55.8%
18/02/2014	76.5%		45.6%			41.5%			59.0%
12/03/2014	86.1%		69.7%			59.7%	51.1%	56.8%	74.7%

Table 17: Goodness of fit of alternative detection function models for predicting probability of detection for rare species.

ID	Model	AIC Term 1	AIC Term 2	AIC	$\Delta$ AIC Term 1	$\Delta$ AIC Term 2	$\Delta$ AIC	Comment
1	1	6901.9	2	6903.9	0.0	0	0.0	Null model
2	length + log(size) + observer + seastate	6681.5	26	6707.5	-220.4	24	-196.4	
3	length <sup>2</sup> + log(size) + observer + seastate	6680.8	26	6706.8	-221.2	24	-197.2	
4	species + log(size) + survey	6649.3	54	6703.3	-252.7	52	-200.7	Base model
5	behaviour + log(size) + observer + seastate	6654.8	26	6680.8	-247.2	24	-223.2	
6	length + behaviour + log(size) + observer + seastate	6648.0	28	6676.0	-254.0	26	-228.0	
7	length <sup>2</sup> + behaviour + log(size) + observer + seastate	6647.4	28	6675.4	-254.5	26	-228.5	Rare species model
8	species + log(size) + observer + seastate	6635.0	40	6675.0	-266.9	38	-228.9	Common species model

Table 18: Coefficients for rare species detection function model

<b>Coefficient</b>	<b>Estimate</b>	<b>SE</b>	<b>CV</b>
Intercept (behaviour: Diving; observer 1)	5.28	0.10	2%
length <sup>2</sup>	0.58	0.24	42%
behaviour : Surface/Aerial Feeder	0.44	0.08	17%
log(size)	0.20	0.04	21%
observer 2	-0.44	0.09	21%
observer 3	-0.38	0.13	35%
observer 4	-0.29	0.09	29%
observer 5	-0.22	0.31	141%
observer 6	-0.02	0.29	1576%
observer 7	-0.10	0.09	90%
observer 8	-0.46	0.08	17%
observer 9	-0.19	0.26	140%
observer 10	-0.45	0.42	92%
seastate	-0.18	0.02	14%

Table 19: Average probability of detection of clusters of each species on each survey from the rare species detection function model.

Table 19a Common species (n ≥30)

Survey	Species							
	Fulmar	Gannet	Shag	Great skua	Kittiwake	Guillemot	Razorbill	Puffin
27/03/2012	74.5%		65.8%	64.8%		52.2%	43.6%	42.4%
18/04/2012	61.1%	89.0%	43.9%	76.5%	89.4%	44.6%	48.0%	40.5%
27/05/2012	78.7%	87.8%	59.3%	82.1%	79.4%	59.2%	54.8%	55.0%
30/06/2012	52.4%	69.1%		48.7%	79.0%	33.0%	35.6%	31.4%
20/07/2012	76.3%	90.7%	78.1%	82.2%		58.0%	56.9%	48.7%
20/08/2012	62.6%	79.6%		67.4%			37.1%	39.0%
11/12/2012	84.4%	79.5%	47.1%			41.2%		
17/02/2013	68.3%		52.6%		60.2%	38.7%	40.8%	
04/03/2013	66.7%		54.9%		59.6%	42.6%		
05/03/2013	67.9%		55.0%			44.3%	46.9%	40.9%
30/03/2013	74.0%		62.6%			44.9%	45.1%	43.0%
03/06/2013	66.2%	85.1%	85.8%	72.9%	80.1%	49.8%	50.6%	47.5%
25/06/2013	69.0%		68.0%	68.3%	90.1%	46.5%	54.3%	43.6%
10/07/2013	74.4%		68.8%	76.1%	79.7%	53.8%	53.5%	47.8%
09/09/2013	70.6%	85.8%	58.2%	79.4%		43.9%	40.8%	45.2%
23/10/2013		69.1%	50.6%		49.8%	35.2%	36.2%	38.8%
18/02/2014	65.6%		47.7%			38.2%		
12/03/2014	78.6%		66.3%			54.0%	60.6%	53.2%

Table 19b Rare species (n&lt;30)

Survey	Species												
	Red-throated diver	Black-throated diver	Great northern diver	Manx shearwater	Storm petrel	Cormorant	Arctic skua	Common gull	Herring gull	Great black-backed gull	Arctic tern	Black guillemot	Little Auk
27/03/2012	59.4%									74.6%			
18/04/2012												37.5%	
27/05/2012							75.8%				82.5%	54.0%	
30/06/2012				52.7%									
20/07/2012							79.8%						
20/08/2012					57.0%								
11/12/2012								61.3%	73.5%	77.0%			
17/02/2013									50.9%	69.8%		32.4%	
04/03/2013										69.8%			
05/03/2013										79.7%		38.9%	40.4%
30/03/2013										70.2%		58.3%	
03/06/2013												53.6%	
25/06/2013			62.4%					73.9%					
10/07/2013												46.9%	
09/09/2013												46.9%	
23/10/2013						57.6%							
18/02/2014												32.7%	47.3%
12/03/2014		59.4%							78.3%	79.3%			

Table 20: Average probability of detection of individuals of each species on each survey from the rare species detection function model.

Table 20a Common species (n ≥30)

Survey	Species							
	Fulmar	Gannet	Shag	Great skua	Kittiwake	Guillemot	Razorbill	Puffin
27/03/2012	84.8%		66.6%	64.8%		55.7%	51.4%	44.7%
18/04/2012	69.7%	89.0%	47.8%	76.5%	94.6%	45.9%	49.2%	43.1%
27/05/2012	86.4%	90.7%	59.3%	85.5%	81.5%	68.9%	57.6%	61.1%
30/06/2012	59.0%	69.1%		48.7%	82.0%	38.0%	45.1%	32.2%
20/07/2012	77.2%	90.7%	80.0%	82.2%		62.8%	61.9%	51.9%
20/08/2012	66.5%	81.3%		67.4%			37.1%	39.0%
11/12/2012	84.4%	79.5%	45.5%			41.2%		
17/02/2013	82.6%		54.5%		60.2%	39.2%	40.8%	
04/03/2013	76.7%		57.2%		59.6%	44.2%		
05/03/2013	80.4%		58.0%			50.7%	49.4%	40.9%
30/03/2013	84.6%		61.6%			45.0%	45.1%	43.0%
03/06/2013	68.3%	85.1%	85.8%	72.9%	84.5%	54.4%	56.8%	48.8%
25/06/2013	72.1%		69.0%	68.3%	93.1%	58.0%	74.9%	45.0%
10/07/2013	76.1%		71.4%	76.3%	79.7%	56.2%	58.1%	48.4%
09/09/2013	78.3%	86.3%	58.2%	80.1%		46.4%	40.8%	45.2%
23/10/2013		74.7%	50.6%		58.2%	37.5%	38.3%	38.8%
18/02/2014	75.3%		47.7%			39.6%		
12/03/2014	84.3%		68.7%			54.0%	60.6%	53.2%

Table 20b Rare species (n<30)

Survey	Species												
	Red-throated diver	Black-throated diver	Great northern diver	Manx shearwater	Storm petrel	Cormorant	Arctic skua	Common gull	Herring gull	Great black-backed gull	Arctic tern	Black guillemot	Little Auk
27/03/2012	59.4%									74.6%			
18/04/2012												37.5%	
27/05/2012							75.5%				82.5%	55.1%	
30/06/2012				52.7%									
20/07/2012							79.8%						
20/08/2012					57.0%								
11/12/2012								61.3%	73.5%	77.0%			
17/02/2013									50.9%	69.8%		32.4%	
04/03/2013										69.8%			
05/03/2013										79.7%		38.9%	40.4%
30/03/2013										70.2%		58.3%	
03/06/2013												53.6%	
25/06/2013			62.4%					73.9%					
10/07/2013												46.9%	
09/09/2013												46.9%	
23/10/2013						57.6%							
18/02/2014												32.7%	47.3%
12/03/2014		59.4%							79.1%	79.3%			



Table 21: Species-specific seasons used to summarise abundance data

Species	Season	From	To
Fulmar	Breeding	May	September
	Non-breeding	October	April
Manx shearwater	Breeding	Mid-March	September
Storm petrel	Breeding	May	October
Gannet	Breeding	Mid-March	September
	Non-breeding	October	Mid-March
Shag	Breeding	March	August
	Non-breeding	September	February
Arctic skua	Breeding	April	August
Great skua	Breeding	April	September
Common Gull	Breeding	April	August
	Non-breeding	September	March
Herring Gull	Breeding	April	August
	Non-breeding	September	March
Great black-backed gull	Breeding	April	August
	Non-breeding	September	March
Kittiwake	Breeding	Mid-March	Mid-August
	Non-breeding	Mid-August	Mid-March
Arctic tern	Breeding	May	August
Guillemot	Colony attendance	March	July
	Chicks on sea/Moulting	August	August
	Non-breeding	September	February
Razorbill	Colony attendance	April	July
	Chicks on sea/Moulting	August	August
	Non-breeding	September	March
Black guillemot	Breeding	April	August
	Non-breeding	September	March
Puffin	Breeding	April	Mid-August
	Non-breeding	Mid-August	March

Months are inclusive. For example, a May to September season would extend from the 1<sup>st</sup> May to the 30<sup>th</sup> September inclusive. For most months of the year, with 30 - 31 days, we define mid-month as midnight on the 15<sup>th</sup>, whilst for February, with 28-29 days, we define mid-month as midnight on the 14<sup>th</sup>. Therefore, a season from Mid-March to Mid-August, for example, would extend from 16<sup>th</sup> March to the 15<sup>th</sup> August inclusive.

Table 22: Details of each the strata used in calculating abundance and density estimates, including area and expected survey effort and numbers of transects expected to be covered if full coverage achieved.

Strata	Area (km <sup>2</sup> )	Expected Number of transects covered	Expected survey effort (km)
Development Area	11.1	5	8.1
Development Area with 1 km buffer	32.1	6	21.5
Development Area with 2 km buffer	52.9	7	37.1
Development Area with 3 km buffer	77.2	9	55.6
Survey Area	112.9	11	79.5

Table 23: The probability that no birds will be seen within the Development Area if it is assumed that the encounter rate (sightings/km) is the same as across whole study area, but sightings follow a Poisson Distribution. For just those species/season combinations where no birds recorded in the Development Area, but are recorded within the Development with a 1 km buffer, and thus all larger strata.

Subset	Species	Season	Figure in Appendix I	Whole Survey Area				Development Area		
				No of surveys	Total Survey effort (km)	Number of sightings	Encounter Rate (sightings/km)	Total Survey effort (km)	Expected number of sightings	Probability of no sightings
Birds on the water	Gannet	Non-breeding	1.6	6	341.9	5	0.015	43.4	0.64	53%
	Shag	Non-breeding	1.8	5	298.8	15	0.050	36.2	1.82	16%
	Herring gull	Non-breeding	1.13	8	496.7	4	0.008	61.6	0.50	61%
	GBB gull	Non-breeding	1.15	8	496.7	8	0.016	61.6	0.99	37%
	Kittiwake	Breeding	1.16	9	710.2	27	0.038	78.5	2.98	5%
	Razorbill	Non-breeding	1.23	8	496.7	16	0.032	61.6	1.98	14%
	Black guillemot	Breeding	1.24	8	623.9	6	0.010	67.2	0.65	52%
	Black guillemot	Non-breeding	1.25	8	496.7	6	0.012	61.6	0.74	48%
	Puffin	Non-breeding	1.27	9	567.5	30	0.053	67.5	3.57	3%
Birds in flight	Shag	Breeding	1.7	11	858.8	7	0.008	93.6	0.76	47%
	Arctic skua	Breeding	1.9	8	623.9	5	0.008	67.2	0.54	58%
	Common gull	Breeding	1.11	8	623.9	4	0.006	67.2	0.43	65%
	Common gull	Non-breeding	1.12	8	496.7	6	0.012	61.6	0.74	48%
	GBB gull	Breeding	1.14	8	623.9	3	0.005	67.2	0.32	72%
	Arctic tern	Breeding	1.18	6	474.3	10	0.021	53.3	1.12	33%
	Black guillemot	Breeding	1.24	8	623.9	2	0.003	67.2	0.22	81%
Combined	Arctic skua	Breeding	1.9	8	623.9	8	0.013	67.2	0.86	42%
	Common gull	Breeding	1.11	8	623.9	5	0.008	67.2	0.54	58%
	Common gull	Non-breeding	1.12	8	496.7	7	0.014	61.6	0.87	42%
	Herring gull	Non-breeding	1.13	8	496.7	6	0.012	61.6	0.74	48%
	GBB gull	Breeding	1.15	8	623.9	3	0.005	67.2	0.32	72%
	Arctic tern	Breeding	1.18	6	474.3	11	0.023	53.3	1.24	29%
	Black guillemot	Breeding	1.24	8	623.9	8	0.013	67.2	0.86	42%
	Black guillemot	Non-breeding	1.25	8	496.7	7	0.014	61.6	0.87	42%
Puffin	Non-breeding	1.27	9	567.5	32	0.056	67.5	3.81	2%	

Table 24: For the 12 complete surveys, for each species the number of times the density estimate for the eastern transects was greater than that for the western transects, and vice versa, and the probability of this or a more extreme result (based on a two sided binomial test) if there was no difference between the two sets of transects in the probability of the highest density being recorded there.

Table 24a: Both in flight and on the water combined

Species	Number of complete surveys			p
	Density higher in East	Density higher in West	Density same in both areas (no birds)	
Red-throated diver	1	0	11	100.0%
Great northern diver	1	0	11	100.0%
Fulmar	5	7	0	77.4%
Sooty shearwater	0	1	11	100.0%
Manx shearwater	1	2	9	100.0%
Storm petrel	2	2	8	100.0%
Gannet	4	8	0	38.8%
Shag	10	1	1	1.2%
Arctic skua	3	2	7	100.0%
Great skua	2	7	3	18.0%
Black-headed gull	1	0	11	100.0%
Common gull	5	0	7	6.3%
Herring gull	2	1	9	100.0%
Great black-backed gull	4	2	6	68.8%
Kittiwake	6	6	0	100.0%
Common tern	1	0	11	100.0%
Arctic tern	4	1	7	37.5%
Guillemot	7	5	0	77.4%
Razorbill	5	7	0	77.4%
Black guillemot	10	0	2	0.2%
Little Auk	1	0	11	100.0%
Puffin	6	5	1	100.0%

Table 24b: Birds on the water

Species	Number of complete surveys			p
	Density higher in East	Density higher in West	Density same in both areas (no birds)	
Red-throated diver	1	0	11	100.0%
Great northern diver	1	0	11	100.0%
Fulmar	4	8	0	38.8%
Manx shearwater	0	1	11	100.0%
Gannet	2	4	6	68.8%
Shag	10	1	1	1.2%
Arctic skua	1	1	10	100.0%
Great skua	4	5	3	100.0%
Common gull	1	0	11	100.0%
Herring gull	1	0	11	100.0%
Great black-backed gull	1	3	8	62.5%
Kittiwake	4	3	5	100.0%
Arctic tern	0	1	11	100.0%
Guillemot	7	5	0	77.4%
Razorbill	3	9	0	14.6%
Black guillemot	8	0	4	0.8%
Little Auk	1	0	11	100.0%
Puffin	6	5	1	100.0%

Table 24c: Birds in flight

Species	Number of complete surveys			p
	Density higher in East	Density higher in West	Density same in both areas (no birds)	
Fulmar	4	8	0	38.8%
Sooty shearwater	0	1	11	100.0%
Manx shearwater	2	1	9	100.0%
Storm petrel	2	2	8	100.0%
Gannet	5	7	0	77.4%
Shag	5	1	6	21.9%
Arctic skua	2	1	9	100.0%
Great skua	3	5	4	72.7%
Black-headed gull	1	0	11	100.0%
Common gull	5	0	7	6.3%
Herring gull	1	1	10	100.0%
Great black-backed gull	3	0	9	25.0%
Kittiwake	6	6	0	100.0%
Common tern	1	0	11	100.0%
Arctic tern	4	1	7	37.5%
Guillemot	6	5	1	100.0%
Razorbill	6	6	0	100.0%
Black guillemot	2	0	10	50.0%
Puffin	5	4	3	100.0%