Comparison of visual and digital aerial survey results of avian abundance for Round 3, Norfolk Region

Eric A Rexstad
University of St Andrews

COWRIE STANDBIRD-09

Comparison of visual and digital aerial survey results of avian abundance for Round 3, Norfolk Region

Louise Burt, Eric Rexstad and Stephen Buckland
Research Unit for Wildlife Population Assessment
Centre for Research into Ecological and Environmental Modelling
University of St. Andrews

November 2009

This report has been commissioned by COWRIE Ltd

Published by COWRIE Ltd.

ISBN: 978-0-9561404-6-3

Preferred way to cite this report:
Burt, M.L., E. Rexstad and S.T. Buckland. **Comparison of visual and digital aerial survey results of avian abundance for Round 3, Norfolk Region.** COWRIE Ltd, 2009.

Copies available from:
www.offshorewind.co.uk
E-mail: cowrie@offshorewind.co.uk

Contact details:

Louise Burt (louise@mcs.st-and.ac.uk) 01334 461805
Eric Rexstad (ericr@mcs.st-and.ac.uk) 01334 461833
Stephen Buckland (steve@mcs.st-and.ac.uk) 01334 461841
Centre for Research into Ecological and Environmental Modelling
The Observatory, Buchanan Gardens
University of St. Andrews
St. Andrews KY16 9LZ

Table of Contents

List of Tables and Figures

Executive Summary

This report describes the analysis of two aerial surveys in the Norfolk Round 3 Zone 5 potential offshore wind farm area in the spring of 2009. One survey employed on-board observers and subdivided the study region into 8 blocks or strata. These strata were surveyed over 5 days of surveying during the period 5 March-20 April. A second survey employed digital video technology to collect animal detections. This survey was completed over 3 days.

Abundance estimates were computed using conventional distance sampling estimators as well as density surface models for the visual surveys. Strip transect methods as well as density surface models were used to compute abundance estimates for the digital video surveys. Rather than producing estimates for each species encountered during the surveys, species were grouped into 'divers', 'gulls', and 'seabird' categories. Hence six estimates for each strata were produced for the visual survey data, and five abundance estimates were produced for the digital video surveys (too few divers were detected to produce estimates using density surface modelling).

It was not feasible to estimate numbers of each species group in the entire region, because the surveys were conducted over a protracted survey period, during which bird numbers and distributions can be expected to vary substantially. We therefore used smaller areas corresponding to the strata (blocks) defined in the design of the visual surveys as the basic unit of comparison. As a result, the digital survey effort was broken into smaller blocks than in the original design of this survey. Likewise, the comparison of estimates produced by the two survey methods could only be on a single stratum where surveys by the two teams were conducted at roughly the same time. The correspondence of abundance estimates for three species groups in this stratum was reasonable for gulls, but confidence intervals for the two surveys do not overlap for the 'seabird' group.

We measured precision of the visual and digital surveys in two ways. First we selected a stratum (NS1) where the data for the two types of surveys were most comparable. But to approximate the precision for the entire study region, we needed to sum estimates over all strata, even though the survey was carried out over a period >5 weeks. We then needed to further approximate the precision over the entire study region for the digital survey. This involved extrapolating the increase in precision observed for visual surveys from the stratum to the study region to the digital surveys. We urge caution in interpreting the measures of precision at the level of the entire survey region.

Precision of the survey method/analysis method/species group fell broadly into three categories. Conventional distance sampling estimates from visual surveys were most precise with coefficients of variation (CVs) on the order of 0.10. Density surface model estimates of visual surveys of gulls and digital surveys of seabirds along with strip transect digital survey estimates of seabirds produced CVs on the order of 0.30, and all other survey/analysis/species combinations had CVs in the range of 0.45-0.66. Survey/analysis methods that have CVs of 0.30 require an annual change (either positive or negative) of around 11% to ensure that the power to detect this trend within 10 years of annual surveys is to be 0.8. These are quite large absolute changes to the study populations, making those survey/analysis methods quite poor at detecting change in avian populations.

Because of the lack of coordination during the conduct of the surveys, this post-hoc comparison is uninformative from the perspective of comparing survey methods. Movements over the protracted survey period, both within the study region and in or out of it, render dubious any exercise in summing estimates over sub-regions, and in comparing sub-region estimates corresponding to dates that differ appreciably. However, we do have some confidence in the estimates of precision associated with each survey method. Based on the data after filtering and matching from these surveys in the Norfolk Round 3 Zone 5, sufficient power is assured only for avian data derived from visual surveys.

We suspect that both data acquisition methods and survey design can be improved to enhance the precision of abundance estimates using digital techniques. Encounter rate variability

(differences in the number of animals detected between transects) drives much of the uncertainty associated with these surveys, and careful attention to survey design can reduce this source of variation that results in these high measures of uncertainty, and consequent poor power to detect change.

Glossary

α	Probability of a Type I error
β	Probability of a Type II error
Power	1 - β
Type I error	Incorrectly concluding that a trend has occurred
Type II error	Incorrectly concluding that no trend has occurred
Two-tailed test	assessing whether there has been a change (either increase or decrease) in the trend
One-tailed test	assessing whether there has been either an increase or a decrease in the trend

Acronyms

Esw – effective strip width

CDS – conventional distance sampling

DSM – density surface model

Units

Length in kilometres

Perpendicular distances in metres

Density in birds per km^2

1. Introduction

This report describes the analysis of the data collected during aerial surveys of seabirds in a region of the North Sea offshore from East Anglia (zone 5). Conventional distance sampling (CDS) methodology (Buckland *et al*., 2001) and density surface modelling (Hedley *et al*., 2004) were used to estimate density and abundance of three groups of birds: divers, gulls and seabirds. (No sea ducks were sighted during these surveys, which were conducted in March and April.)

Given the two survey methods, and two abundance estimation methods associated with each survey method, precision with which abundance of various seabird species groups could be estimated was of interest. These measures of precision were further translated into the ability of these methods to detect changes in abundances over time. Hence the power of methods for trend detection was assessed.

A trend can be identified when a regression of abundance estimates on year has a slope that is significantly different from zero. A statistical test can be used to identify whether or not the data are consistent with a null hypothesis of no trend. We want the test to reject the null hypothesis of no trend when a trend exists. The probability that the test does this is called the power of the test. In the context of determining changes in avian abundance, we are interested in assessing uncertainty in abundance estimated at multiple points in time compared to the magnitude of overall change in abundance. The power of the tests associated with the different survey and analysis methods can be calculated and compared using the methodology of Gerrodette (1987).

2. Methods

2.1 Data collection

2.1.1 Visual surveys

Because the region of interest could not be surveyed in a single day, it was divided into seven survey blocks or strata (Figure 1), and a small aircraft flew a series of transects aligned north-south and spaced 2km apart within each block (see Taylor-Jones, 2009). Flights took place between 5 March 2009 and 30 April 2009 (Table 2). The plane flew at an altitude of 76m and at a speed of 200km/h approximately. Observers searched both sides of the plane and for each flock of birds detected recorded the species, number in the flock, behaviour and the perpendicular distance to the transect. These distances were allocated to four distance bands covering a region from 44m to 1000m either side of the plane.

2.1.2 Digital video surveys

The aircraft flew a series of transects spaced approximately 2-3km apart. Four cameras recorded along each transect; each camera recording a strip 50m wide. Any sightings recorded were categorised by species (with associated confidence of category into which the sighting was placed) and behaviour. If birds were very close to each other, a count of more than one bird might be recorded at a single location, but no attempt was made to define bird flocks. In analysis therefore, we made no attempt to estimate or model flock size. See Thaxter and Burton (2009) for further discussion of digital survey methodology.

2.1.3 Preparation of data for analysis

Surveys employing digital video were flown on 3 days (18 March, 20 April and 22 April), and were not flown according to the stratification pattern used in the visual surveys shown in Figure 1. Therefore, the data provider for the digital video data graciously segmented their data to conform to the regions surveyed by the visual survey providers. The timings of the digital and visual surveys were not scheduled to coincide, so data from only 4 of the survey blocks were included for comparison in this report (Figure 2). Date of visual surveys for NS2 and NS3 were a fortnight prior to the first digital video survey, and were hence excluded. Technical difficulties challenged clipping of digital survey data corresponding with the boundaries of NS5, and the further subdivision of NS7 into two sections made the abundance estimate derived for that stratum uninformative.

2.2 Data analysis

2.2.1 Conventional distance sampling methods

Visual survey methods require the fitting of detection functions to the observed detections to account for birds not detected during the survey (Buckland *et al.*, 2001). The digital video surveys generate digital images which can be examined in detail, so that it is reasonable to assume that no birds sitting on the surface of the water or flying within the surveyed area were undetected.

2.2.1.1 Visual surveys

Conventional distance sampling methodology (Buckland *et al*., 2001) implemented in the program Distance (Thomas *et al*., 2009) will be used to estimate abundance (*N*) in each block (*b*) as follows

$$\hat{N}_b = A_b \frac{n_b}{2L_b\hat{\mu}}\hat{E}[s]$$

where A_b is size of the survey block, n_b is the number of detected flocks, L_b is the length of transects searched, $\hat{\mu}$ is the estimate of the effective search half-width (ESW, the distance from the transect line beyond which as many birds are detected as are missed at distances smaller than ESW), and $\hat{E}[s]$ is the estimate of the mean flock size. The ESW is obtained from a detection function model fitted to the distribution of perpendicular distances. The expected flock size is obtained from a regression of probability of detection (obtained from the detection function) against the logarithm of flock size. This takes into account the difficulty of seeing single birds and small flocks further away from the plane.

2.2.1.2 Digital video surveys

Because all birds within the strip are detected, a simple strip transect estimator can be used to estimate abundance in each survey block *b* as follows

$$\hat{N}_b = A_b \frac{n_b}{2wL_b}$$

where A_b is size of the survey block, *w* is the strip half-width, n_b is the number of detected birds, L_b is the length of transects searched.

2.2.2 Density surface modelling

2.2.2.1 Visual surveys

Trend in the spatial distribution of the species groups were modelled using the ‘count’ method of Hedley and Buckland (2004). In this approach, the transect lines are divided into small segments (of length l_i) and the response variable is the number of birds counted within the segment, taking into account the probability of detection:

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{s_j}{\hat{p}}$$

where $\hat{N}_i$ is the number of birds estimated to be in the segment *i*, s_j is the recorded flock size for flock *j* and n_i is the number of flocks recorded in segment *i*. The parameter $\hat{p}$ is the probability of detection and is obtained from $\hat{p} = \hat{\mu}/w$ where *w* is the truncation distance beyond which sightings are excluded. See previous section for a description of $\hat{\mu}$. The response variable, $\hat{N}_i$, is modelled as a function of covariates with a generalised additive model (GAM) with the general formulation

$$E\left[\hat{N}_i\right] = \exp\left[\log(a_i) + \beta_0 + \sum_{k=1}^{K} f_k(z_{ik})\right]$$

The term log(a_i) is an offset (a term with known regression coefficient) that corresponds to the area of the segment ($a_i = 2wl_i$), β_0 is the intercept and f_k are smooth functions of the K covariates. This formulation assumes a logarithmic link function and an overdispersed Poisson distribution for the error distribution. Surveys of animals (particularly those that exhibit clumped distributions) tend to have portions of the survey region where there are no detections and others where there are many. This leads to overdispersion in the data, which may be modelled using a quasi-Poisson error distribution. The logarithm is the canonical link function for Poisson or quasi-Poisson error distribution families.

Having obtained a fitted model from the data, density can be estimated throughout the region of interest; abundance is obtained by integrating under this surface. The variance of the abundance is obtained by a bootstrap – a data-based simulation method. A sample is drawn from the original data by sampling lines at random and with replacement and abundance is estimated from this bootstrap sample; the process is repeated a large number of times and the empirical variance is calculated from the distribution of the bootstrap abundance estimates.

2.2.2.2 Digital video surveys

Methods of fitting GAMS to the digital data were identical to that described in section 2.2.2.1, with the exception that the number of birds in each segment did not need to be estimated, and was merely the sum of all birds counted in the segment

$$N_i = \sum_{j=1}^{n_i} s_j$$

where N_i is the number of birds in segment i, s_j is the recorded flock size for flock j and n_i is the number of flocks recorded in segment i.

2.3 Power calculations

Detection of a trend requires five parameters shown in Table 2 to be specified (Gerrodette, 1987). Given any four of these parameters, the fifth can be found. Here we are interested in calculating the power of the test (1-β). Therefore, we need to specify the four other parameters; the number of time steps or annual surveys, the rate of change of the quantity being measured, the coefficient of variation (CV) of abundance estimates and the probability of Type I error (rejection of a null hypothesis that is true).

We take the estimated CV for each of the survey and analysis methods, and feed these into the power analysis according to this formula

$$ncp = \frac{\log(1+r)}{\sqrt{\dfrac{\log(CV(\hat{N})^2+1)}{t(t-1)(t+1)/12}}}$$

$$1-\beta = 1-\Phi(z_{1-\alpha/2} - ncp) + \Phi(z_{\alpha/2} - ncp)$$

where

r is population rate of change,

CV(N) is coefficient of variation in abundance, t is number of samples,

ncp is a computed non-centrality parameter associated with the standard normal distribution (Φ), and

z_α is the α-level value from a standard normal distribution (a t-distribution was shown in Gerrodette's (1991) paper to be preferable, and that was used in our calculations).

We will present the simple situation where we wish to know the probability of detecting a change in avian abundance when a change has actually occurred (power) using the measures of precision for NS1. This is the stratum where the surveys were conducted most close in time, and were therefore most directly comparable. To approximate the power associated with detecting change within the entire Norfolk study region, we make several approximations. First, we sum the estimates from the visual surveys across all strata, even though this results in meaningless abundance estimates because of the protracted duration of the survey. The approximation is compounded for the digital surveys because abundance estimates were not computed for the entire study region. Therefore, the proportional decrease in CV for region-wide estimates from digital surveys were assumed to be equivalent to the proportional decrease in CV between visual estimates at the stratum level and visual estimates at the region level. This amounts to a decrease by 60% in the visual strip transect estimates of divers to a 10% reduction in the strip transect estimates for seabirds.

Our remit of comparing the visual and digitally collected data led us to use the visual survey data blocks as the 'least common denominator' for our comparisons, and the digital video data were altered to suit the visual survey design. Were we to have used the digital video data in their entirety to compute an estimate of abundance and its precision for the entire Norfolk study region, the coefficient of variation for the estimate for the study region would have been smaller than that produced by the method described in the previous paragraph. This was not done because the digital video data had been adjusted to conform to the visual survey design strata by the data provider.

Power calculations follow the methods of Gerrodette (1991) and assume a linear annual change in abundance, surveys equally-spaced in time, and use a t-distribution. The change in CV as abundance changes is assumed to be proportional to the square root of abundance (see Gerrodette, 1987).

3. Results

3.1 Search effort and number of sightings

3.1.1 Visual surveys

The blocks were surveyed on four different dates in 2009 (Table 2) and since birds can move a substantial amount between the different days, estimates are obtained separately for each survey block. Block NS7 was surveyed on two different dates and so this block has been divided into two (called A and B) and estimates obtained separately for each date.

The distance flown along each transect was calculated using the start and end positions of each transect. A total of 3,264km of transects were covered on search effort.

Seabirds (defined as auks, fulmars and gannets, see Table 3) were the most frequently sighted species group with 1,956 detections and divers were the least sighted group (402 detections); 992 detections of gulls were used in the analysis of abundance.

Maps of search effort and locations of sightings are shown in Figure 3 and a summary of the data collected is given in Table 2.

3.1.2 Digital video surveys

Data for four blocks are available; blocks NS1, NS2 and NS4 were surveyed on 22 April, 2009 and block NS7 was surveyed on 18 March, 2009. No areas covered during survey effort on 20 April 2009 were used in this report. A total of 1,679km of transect were covered. To increase sample size, species were categorised into the species groups shown in Table 3. Sightings with a missing species were deleted, along with sightings that did not fall along transects; however, all levels of confidence of species identification have been included. Gulls were the most frequently sighted species group with 119 flocks detected; divers were the least sighted group (9 flocks); 93 flocks of seabirds were detected (Table 4).

3.2 Distance sampling or strip transect estimates

3.2.1 Visual surveys

Detection function

The region from directly under the plane out to 44m could not be seen by the observers. Therefore, the first distance bin necessarily starts at 44m, hence 44m was subtracted from all perpendicular distances and so the cutpoints for the bins became 0, 119, 238, 382 and 956m. Very few sightings were detected in the furthest distance band and so to avoid a long tail in the fitted model the data were truncated at 382m leaving 402 detections of divers, 992 gull detections and 1956 seabird detections.

Detection functions were fitted for each species group and both the half-normal and hazard rate forms of the detection function were considered. Although abundance estimates were required for each block, there were not enough sightings in some blocks to be able to fit separate detection functions and so a single detection function with sightings pooled over all blocks was estimated. The half-normal form had a smaller AIC, but the difference between the half normal fit and hazard rate fit was small (ΔAIC <2.1 in all cases). It was felt that the hazard rate form was more appropriate for this data because it forced the function to have a wider 'shoulder,' a desirable consideration for detection functions (Buckland *et al,.* 2001:42). The two-parameter hazard rate model will also be conservative because it will produce abundance estimates with larger variances as there is an additional parameter to be estimated. The fitted models are shown in Figure 4 and the ESW are given in Table 5.

Flock size

The vast majority of detections were of single birds, however flocks of up to 15 birds were detected for divers and seabirds and of up to 150 birds for gulls. As with the detection function, a single estimate of flock size for each species group was obtained for the region of interest using the default regression method of Distance to correct for size-biased detections. Larger flocks tend to be more detectable, so that the mean size of detected flocks tends to be larger than the mean size of all flocks in the survey region. By regressing log flock size on estimated probability of detection (as a function of distance from the line), we estimate mean flock size when detection is certain (on the line), as described by Buckland et al. (2001:73-75). Estimates of flock size are given in Table 5 and estimates of bird abundance are given in Table 6.

3.2.2 Digital video surveys

Encounter rate was obtained in each block. Counts of flock size were converted into multiple detections of single individuals as how flocks were determined was uncertain. The strip half-width was taken to be 100m. Density and abundance estimates are given in Table 7.

3.3 Density surface modelling estimates

3.3.1 Visual surveys

Hedley and Buckland (2004) recommended that segments are approximately square and so with a truncation distance of 382m, transects were divided into segments of approximately 1km resulting in 3,266 segments. The number of segments that contained sightings for each species group is shown in Table 8.

The detection function model fitted in the CDS analysis was used in the DSM analysis; however, the detection probabilities shown in Table 8 differ very slightly from those in the CDS (Table 5) because the number of sightings used was slightly different: some sightings has been allocated to transects but were not allocated to segments because they were before the start time or after the end time of the transects. Figure 5 (left column) shows the distribution of sightings by segment taking into account the detection probability. Note the large estimated number of gulls in some segments were due to large flock sizes (55 and 80 in NS3 and 100 and 150 in NS6).

The explanatory variables available for inclusion in the density surface model were the easting (*east*) and northing (*north*) of the midpoint of the segment as both one-dimensional functions and as a two dimensional function; survey block (*block*) was also included as a factor variable. *Month* was not included as a covariate, because that would have been completely confounded with *block*. Akaike's Information criterion was used to choose between the different models. The final models are shown in Table 8.

The estimated density surfaces for each species group are shown in Figure 5 (right column). The abundances are shown in Table 9.

3.3.2 Digital video surveys

There were too few flocks of divers detected to be able to estimate a density surface. The other species groups were more numerous but still sparse in blocks NS1 and NS4. Many segments with no sightings can cause problems in model fitting and so to avoid too many zero segments, transects were divided into segments of approximately 5km in length. This resulted in 338 segments. Detections by species group along the track lines are shown in Figure 6.

The explanatory variables available for inclusion in the model were the location of the segment (in terms of latitude and longitude), survey block number and month of survey (which groups blocks NS1, NS2 and NS4 together). The location variables were fitted as both one-dimensional smooth functions and as a two-dimensional function. The latter two variables were fitted as factor variables. The final models were chosen on the basis of the Generalised Cross Validation score and are shown in Table 10. The abundance estimates are shown in Table 11. Figure 7 shows the spatial depiction of the estimates of density from the fitted density surface models for gulls and seabirds.

To overcome problems in fitting the models to bootstrap samples, the degrees of freedom allowed for the smooth function were fixed at the number used in the original model (see Table 10). Despite this, occasionally a bootstrap sample generated was so extreme that it produced unrealistically high abundance estimates. Such values can distort the variance of the distribution and so outliers were identified using the method of Hoaglin *et al*. (1983) and removed before estimating the variance.

3.4 Trend detection

The relative precision of abundance estimates for the combinations of survey method, analysis type, seabird species group and spatial extent is shown in Table 12. The power of the various scenarios is of interest but as well as the four parameters already described, two other options need to be specified. These are whether the test is one- or two-tailed and whether the rate of change is the overall change between the start and end of the study or is incremental each year.

Figure 8 shows a set of power curves derived from the assumptions presented in section 2.3 related to stratum NS1. This is the stratum of the Norfolk study region wherein 468km of transect was flown on 20 March for the visual survey, and 414km of transect was flown on 18 March for the digital video survey. There are many ways of depicting power; in Figure 8 we show the power to detect annual change in the population of interest if the population is changing annually at the rate depicted on the x-axis. We arbitrarily show the power to detect the change after a period of 10 years of change has transpired (the rank of the survey designs, species groups and analysis methods would remain unchanged for different length of time transpired). For the species groups recognised in these surveys and for a survey area of the size of NS1 (917km^2) in March, power for comparisons drawn from digital survey data is universally lower than the power for visual survey data.

A similar graph is provided in Figure 9 where power calculations are extrapolated to the entire study region. These calculations are based on tenuous assumptions (i.e., the abundance and precision estimates of visual surveys conducted over a 5-week period are meaningful). This is compounded by estimating that the precision for digital surveys over the entire region would reduce by the same proportion that the visual survey precision decreased. The pattern in this graph is similar to Figure 8, in that power is quite high for several of the visual survey analysis and species groups (conventional distance sampling analysis of all three species groups). The remainder of power curves constitute an interspersion between visual and digital data.

If we were to construct power curves under a different scenario, e.g., 5 years of surveying rather than 10, an annual change of 0.14 could be detected when the CV is 10%, but the annual change would need to be on the order of 0.45 for surveys with a CV of 30% to have a power of 0.8, and 0.75 annual change could be detected with surveys with a CV of 45%.

4. Discussion

4.1 Visual surveys

Seabirds were the most common species group encountered and divers the least common. For all species groups, the encounter rates were very different between the survey blocks, with smaller numbers being detected in blocks NS5 and NS7B. This has resulted in very different abundances between the survey blocks.

Observers appeared to have concentrated on detecting animals in the first distance band (44 – 163m) and so there are many more sightings in this band than in the outer bands. While ensuring that certain probability of detection in the first band is important, this has resulted in rather 'spiked' detection functions. An alternative analysis would be to use only the first distance band and treat it as a strip transect, but this analysis was not conducted given the time available for our comparison.

The CDS and DSM block abundances were very similar for divers and seabirds; the differences in total abundance were less than 3%. The difference in the two estimates for gulls was more substantial, with the DSM total abundance estimate 25% larger than the CDS estimate. This difference was mainly due to differences in blocks NS3 and NS6 and these were the blocks where the large flock sizes had been detected. The DSM approach models individuals (taking into account recorded flock size) whereas in the CDS estimator, an expected flock size pooled over the whole region has been applied to obtain bird abundance.

The CV estimates were substantially lower for the CDS analysis than the DSM analysis for divers and gulls but were very similar for seabirds. This may be related to the patchiness in distribution of divers and gulls (present in < 20% of the surveyed segments), whereas seabirds were present in >30% of surveyed segments (Table 8).

4.2 Digital video surveys

Gulls were the most frequently detected group. Divers were the group least frequently detected and there were too few detected to be able to produce a DSM estimate of abundance. Most birds were detected in block 7 which was surveyed a month earlier than the other blocks. Taking this into account in the strip transect estimates, by estimating a separate flock size for block 7, made little difference to the overall abundance estimates compared to estimating abundance with a single flock size for all blocks. Month was chosen as an explanatory variable in the DSM for gulls but not for seabirds.

The DSM estimates were approximately 9% lower than the strip transect estimates, and the coefficients of variation for the two analytical methods for digital surveys were also fairly close.

4.3 Comparison of point estimates for different survey methods

The comparison of abundance estimates between the visual and digital video survey methods was made difficult by the size of the Norfolk study region, and the subsequent need to subdivide the study region into strata or blocks for the visual surveys (Taylor-Jones, 2009). Even within a stratum, it was not possible to complete a visual survey within a one-day window (i.e., stratum 7) such that the overall estimate of abundance for this stratum is suspect, and is not included in abundance comparison. This made comparison of abundance for the entire study region impossible, and the temporal incongruence of survey effort for the two methods made comparison all the more difficult. The most thorough comparison that can be made is located in Table 13.

For stratum 1, the abundance estimates for gulls derived from data for both methods is quite close (estimates from digital video data were 87% of the estimates from visual survey data from both analytical methods). The seabird estimates for this stratum from the digital video data were less than half the estimates produced from the visual survey for these surveys conducted two days apart. This is in contrast with findings of Rexstad and Buckland (2009) who compared visual and digital estimates of abundance of common scoters in Carmarthen Bay with both surveys conducted on the same day, and found point estimates from digital video surveys to be consistently larger than visual survey point estimates.

Strata 4 and 6 were surveyed by the two methods more than a month apart with the digital video surveys conducted in late April whereas the visual surveys were flown in mid-March, and from the resulting estimates it is apparent that birds had redistributed during that interval such that further comparison of the point estimates is not meaningful.

4.4 Precision comparison for survey and analysis methods

The coefficients of variation (Table 12) associated with the abundance estimates fall broadly into 3 categories. There are CVs in single digits that include conventional distance sampling estimates for all species groups from the visual surveys and density surface modelling for seabirds. The second group of precision measures of approximately 30% were of density surface estimates for gulls from visual surveys and digital video survey data for seabirds using either analysis method. The final group were CVs between 40% and 50% for density surface estimates of divers from visual surveys, and gull estimates from digital surveys for both analysis methods, as well as strip transect estimates for diver detections from digital video data. We have considerably more confidence in the measures of precision derived from stratum NS1 (Table 12 top) than in the measures of precision for either visual or digital methods, derived for the entire study region (Table 12 bottom).

4.5 Points to consider

Encounter rate variability (variability in number of animals encountered between transects) dominates the uncertainty in the estimation of abundance (see Tables 5 and 7). So understanding the mechanisms giving rise to encounter rate variability will aide in our understanding of uncertainty, and hence power of our surveys to detect change. For species with clumped distributions, such as seabirds, the encounter rate variance is far and away the largest source of variability in the estimation process.

The grouping of species into three species groups will have an impact upon the findings of this report. The reason for using species groups is this was how the digital video data were provided to us, which is likely a result of the image reading process. It is likely that combining species into groups could 'homogenize' encounter rate variability if there is spatial segregation between species within the same grouping. Alternatively if the species within groups share habitat affinity, this may magnify the encounter rate variability, having knock-on effects for the CV of the estimated abundances.

When performing density surface modelling, in which transects are broken into segments, we can view encounter rate variability not at the level of the transect, but at the level of the segment. Note in the summary of density surface modelling for visual (Table 8) and digital video (Table 10) the proportion of segments with detections is small (<20%) for visual surveys of divers and gulls and digital surveys of gulls and seabirds. This resulted in high CVs (>30%) for those abundance estimates. However, for visual surveys of seabirds, >33% of surveys had detections, and the associated CV of the abundance estimate was <10%. So, surveys in which a large proportion of segments have detections lead to density surface modelling estimates of abundance with greater precision.

The spectre of encounter rate variance percolates to the power calculations presented in Figures 8 and 9. The survey/analysis/species combinations that have low variability in encounter rates in turn have low CVs in estimates of abundance. These low CVs for abundance estimates result in surveys that have high power to detect change under the circumstances of linear annual change in abundance over a period of 10 years.

5. Conclusion

Despite the best efforts of the data providers and analysts, it was a challenge to reconcile data collected at different times using different sampling schemes and different sampling equipment and protocols. We were more successful in contrasting estimates derived from visual and digital survey methods from a coordinated study conducted in Carmarthen Bay for common scoter (Rexstad and Buckland, 2009). When estimates from two different methods do not agree, and truth is unknown, it is not possible on statistical grounds to say which method (if either) provides reliable estimates.

Combining two methods of data collection, two methods of analysis for each type of data collection, and three species groups leads to 12 estimates of precision (and associated power) to compare. Highest power was achieved for conventional distance sampling analysis of visual survey data for gulls and divers, and density surface modelling analysis of visual survey data for seabirds.

However on the basis of only two comparisons of estimates derived from survey methods, it is not obvious that one method is universally preferred. Continued efforts should be made to incorporate rigorous comparison of survey method estimates under a variety of circumstances from the time the surveys are first proposed and tenders drawn up.

References

Buckland S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. and Thomas, L. (2001). *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford, UK.

Gerrodette, T. (1987). A power analysis for detecting trends. Ecology **68**,1364-1372.

Gerrodette, T. (1991). Models for power of detecting trends: a reply to Link and Hatfield. Ecology **72**,1889-1892.

Hedley, S.L. and Buckland, S.T. (2004). Spatial models for line transect sampling. Journal of Agriculture, Biological and Environmental Statistics. **9**,181-199.

Hoaglin, D., Mosteller, F. and Tukey, J. (1983). Understanding robust and exploratory data analysis. John Wiley and Sons, New York.

Rexstad. E., and Buckland, S. T. (2009). Comparison of aerial survey methods for estimating abundance of common scoters. Final report submitted to Defra, July 2009.

Taylor-Jones, R. (2009). *Aerial surveys of Round 3, zone 5 for waterbirds – final report*. WWT Consulting.

Thaxter, C. B. and Burton, N. H. K. (2009). High Definition Imagery for Surveying Seabirds and Marine Mammals: A Review of Recent Trials and Development of Protocols. Report submitted to COWRIE.

Thomas, L., Laake, J.L., Rexstad, E., Strindberg, S., Marques, F. F. C., Buckland, S. T., Borchers, D. L., Anderson, D. R., Burnham, K. P., Burt, M. L., Hedley, S. L., Pollard, J. H., Bishop, J. R. B. and Marques, T. A. (2009). Distance 6.0. Release 1. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/.

Table 1. Summary of visual survey data for each survey block showing the date on which the block was covered, the size of the survey block, the number of transects (*k*), search effort (*L*) and the number of flocks of each species group seen during search effort. Note that the area of blocks NS7A and B have been estimated from the transect coverage.

Block	Date	Area (km^2)	*k*	*L* (km)	Divers	Gulls	Seabirds
NS1	20 March	916.6	19	467.8	75	180	342
NS2	5 March	968.3	20	480.6	145	218	230
NS3	5 March	934.7	20	462.8	79	124	664
NS4	18 March	997.1	15	491.2	48	174	113
NS5	20 March	1040.7	20	478.3	1	15	97
NS6	18 March	1015.6	21	492.8	23	206	410
NS7A	31 March	*550.0*	9	230.8	26	90	118
NS7B	20 April	*340.0*	11	160.1	6	1	19
Total		6763.0	135	3264.3	403	1008	1993

Table 2. Parameters required in a power analysis to be able to detect a trend. Power is defined as 1- β.

Parameter	Description
n	Number of time steps in study (eq. years)
r	Rate of change of the quantity being measured as either 1. overall change between start and end 2. change between each time step
CV	Coefficient of variation of abundance
α	Probability of a Type I error, usually 0.05
β	Probability of a Type II error

Table 3. Categorisation of species group for both visual and digital video surveys.

Species group	Species category
Diver	Diver
Gull	Black backed gull Gull Herring gull
Seabird	Auk Fulmar Gannet

Table 4. Summary of digital video survey data. Numbering convention between digital video survey and visual survey depicted in left-most column.

Block	Date	Area (km^2)	k	L (km)	Divers	Gulls	Seabirds
1 (NS7)	22 April	698.5	15	340.84	0	1	2
2 (NS6)	22 April	911.2	17	453.52	1	41	33
4 (NS4)	22 April	966.8	21	470.42	0	10	2
7 (NS1)	18 March	929.4	17	413.97	8	67	56
Total		3505.9	70	1678.75	9	119	93

Table 5. Estimates of encounter rate (*n*/*L*), esw ($\hat{\mu}$) and expected flock size ($E[\hat{S}]$) for each species group in the visual surveys. The data have been truncated at 382m. Percentage coefficients of variation are given in parentheses.

Species	Block	*n*/*L* (flocks/km)	$\hat{\mu}$ (m)	$E[\hat{S}]$
Divers	NS1	0.158 (19.5)	152.9 (4.03)	1.34 (2.28)
	NS2	0.302 (17.9)		
	NS3	0.171 (37.7)		
	NS4	0.098 (22.5)		
	NS5	0.002 (100.0)		
	NS6	0.047 (30.0)		
	NS7A	0.113 (19.1)		
	NS7B	0.038 (54.0)		
Gulls	NS1	0.385 (10.1)	192.6 (3.03)	1.53 (2.24)
	NS2	0.439 (16.0)		
	NS3	0.266 (13.2)		
	NS4	0.354 (19.4)		
	NS5	0.031 (35.5)		
	NS6	0.404 (22.0)		
	NS7A	0.386 (21.6)		
	NS7B	0.006 (96.6)		
Seabirds	NS1	0.710 (9.65)	154.04 (3.39)	1.42 (1.07)
	NS2	0.454 (19.1)		
	NS3	1.424 (20.1)		
	NS4	0.222 (16.3)		
	NS5	0.201 (16.2)		
	NS6	0.826 (20.1)		
	NS7A	0.503 (21.6)		
	NS7B	0.119 (35.6)		

Table 6. CDS estimates of flock density ($\hat{D}_c$) and abundance ($\hat{N}_c$) and individual bird density ($\hat{D}$) and abundance ($\hat{N}$) for each species group and survey block for visual survey data.

Species	Block	$\hat{D}_c$ (flocks/km^2)	$\hat{N}_c$	$\hat{D}$ (birds/km^2)	$\hat{N}$
Divers	NS1	0.52 (19.9)	474 (19.9)	0.69 (20.0)	635 (20.0)
	NS2	0.99 (18.4)	955 (18.4)	1.32 (18.5)	1280 (18.5)
	NS3	0.56 (37.9)	522 (37.9)	0.75 (37.9)	699 (37.9)
	NS4	0.32 (22.9)	319 (22.9)	0.43 (23.0)	427 (23.0)
	NS5	0.01 (100.1)	7 (100.1)	0.01 (100.1)	10 (100.1)
	NS6	0.15 (30.3)	155 (30.3)	0.20 (30.3)	208 (30.3)
	NS7A	0.37 (19.5)	203 (19.5)	0.49 (19.6)	272 (19.6)
	NS7B	0.12 (54.2)	42 (54.2)	0.16 (54.2)	56 (54.2)
	Total	0.40 (11.7)	2676 (11.7)	0.53 (11.9)	3586 (11.9)
Gulls	NS1	1.00 (10.6)	916 (10.6)	1.52 (10.8)	1401 (10.8)
	NS2	1.14 (16.3)	1104 (16.3)	1.74 (16.4)	1689 (16.4)
	NS3	0.69 (13.5)	645 (13.5)	1.06 (13.7)	987 (13.7)
	NS4	0.92 (19.6)	917 (19.6)	1.41 (19.7)	1403 (19.7)
	NS5	0.08 (35.6)	85 (35.6)	0.12 (35.7)	130 (35.7)
	NS6	1.05 (22.2)	1065 (22.2)	1.60 (22.3)	1629 (22.3)
	NS7A	1.00 (21.8)	551 (21.8)	1.53 (21.9)	843 (21.9)
	NS7B	0.02 (96.6)	6 (96.6)	0.02 (96.6)	8 (96.6)
	Total	0.78 (7.89)	5287 (7.89)	1.20 (8.20)	8091 (8.20)
Seabirds	NS1	2.30 (10.2)	2111 (10.2)	3.27 (10.3)	2998 (10.3)
	NS2	1.47 (19.4)	1426 (19.4)	2.09 (19.4)	2025 (19.4)
	NS3	4.62 (20.3)	4320 (20.3)	6.56 (20.4)	6135 (20.4)
	NS4	0.72 (16.6)	718 (16.6)	1.02 (16.7)	1020 (16.7)
	NS5	0.65 (16.5)	678 (16.5)	0.93 (16.6)	963 (16.6)
	NS6	2.68 (20.4)	2722 (20.4)	3.81 (20.4)	3866 (20.4)
	NS7A	1.63 (21.8)	897 (21.8)	2.32 (21.9)	1274 (21.9)
	NS7B	0.39 (35.8)	131 (35.8)	0.55 (35.8)	186 (35.8)
	Total	1.92 (9.18)	13005 (9.18)	2.73 (9.25)	18467 (9.25)

Table 7. Strip transect estimates from digital video surveys of encounter rate (n/L), flock density ($\hat{D}_c$) and abundance ($\hat{N}_c$) and individual bird density ($\hat{D}$) and abundance ($\hat{N}$) for each species group and survey block.

Species	Block	n/L (flocks/km)	$\hat{D}$ (birds/km^2)	$\hat{N}$
Divers	1	0		
	2	0.002 (97.7)	0.011 (97.7)	10 (97.7)
	4	0		
	7	0.024 (48.9)	0.121 (48.9)	110 (48.9)
	Total	0.007 (47.6)	0.035 (45.6)	122 (45.6)
Gulls	1	0.003 (97.6)	0.015 (97.7)	10 (97.7)
	2	0.119 (47.3)	0.595 (47.3)	542 (47.3)
	4	0.021 (37.4)	0.106 (37.4)	103 (37.4)
	7	0.263 (63.7)	1.317 (63.7)	1224 (63.7)
	Total	0.104 (44.6)	0.536 (43.7)	1879 (43.7)
Seabirds	1	0.009 (74.1)	0.044 (74.1)	31 (74.1)
	2	0.176 (43.0)	0.882 (43.0)	804 (43.0)
	4	0.004 (65.7)	0.021 (65.7)	21 (65.7)
	7	0.271 (45.1)	1.353 (45.1)	1257 (45.1)
	Total	0.117 (33.6)	0.602 (31.5)	2112 (31.5)

Table 8. Number of segments from visual survey data that contain sightings (out of 3266 segments), the probability of detection for each species group ($\hat{p}$), the covariates used in the final model and the percentage of deviance explained by the model. The term $s(x, y)$ represents a 2D smooth function of variables x and y.

Species group	Number of segments >0	$\hat{p}$ (%CV)	Covariates	%Deviance
Divers	311	0.40 (4.09)	*Block* + *s*(*east*, *north*)	29.8
Gulls	620	0.51 (2.98)	*Block* + *s*(*east*, *north*)	30.8
Seabirds	1116	0.42 (2.86)	*Block* + *s*(*east*, *north*)	35.4

Table 9. Density surface modelling estimates of abundance ($\hat{N}$) for each species group and survey block from the visual survey data.

Species	Block	$\hat{N}$
Divers	NS1	559 (15.0)
	NS2	1366 (18.2)
	NS3	835 (54.4)
	NS4	399 (26.2)
	NS5	7 (98.0)
	NS6	149 (50.6)
	NS7A	304 (29.8)
	NS7B	55 (54.3)
	Total	3674 (40.6)
Gulls	NS1	1347 (21.2)
	NS2	1357 (16.4)
	NS3	2065 (24.4)
	NS4	1691 (21.4)
	NS5	85 (45.1)
	NS6	3521 (88.7)
	NS7A	726 (38.5)
	NS7B	5 (98.1)
	Total	10798 (31.0)
Seabirds	NS1	2898 (9.15)
	NS2	2133 (14.6)
	NS3	5868 (9.58)
	NS4	981 (14.5)
	NS5	809 (14.7)
	NS6	3870 (17.1)
	NS7A	1427 (19.0)
	NS7B	170 (29.7)
	Total	18155 (6.67)

Table 10. Number of segments that contain sightings (out of 338 segments), the covariates included in the final model and the percentage of deviance explained by the model. The term $s(x, d)$ represents a 1D smooth function of variables x with d degrees of freedom and $s(x, y, d)$ represents a 2D smooth function of variables x and y with d degrees of freedom.

Species group	Number of segments >0	Covariates	%Deviance
Gulls	49	$s(lon, lat, 18.5) + month$	45.8
Seabirds	44	$s(lon, 8.5) + s(lat, 5.9)$	58.6

Table 11. Density surface estimates of abundance ($\hat{N}$) for each species group and survey block derived from digital video survey.

Species	Block	$\hat{N}$
Gulls	1	10 (80.9)
	2	441 (36.9)
	4	100 (50.2)
	7	1173 (66.3)
	Total	1724 (49.4)
Seabirds	1	25 (70.1)
	2	660 (13.1)
	4	12 (95.6)
	7	1182 (44.5)
	Total	1879 (29.3)

Table 12 CVs (as a percentage) associated with the overall abundance estimates for the two different survey methods and the two different analysis methods for each species group. A density surface analysis of divers from the digital data was not possible.

Species group	Analysis method			
	Line/strip transect		Density surface modelling	
Stratum NS1				
	Visual	Digital	Visual	Digital
Divers	12	49	41	—
Gulls	11	64	21	66
Seabirds	10	45	9	45
All Norfolk study region				
	Visual	Digital	Visual	Digital
Divers	12	29	41	—
Gulls	8	47	31	51
Seabirds	9	41	7	35

Table 13. Portions of the Norfolk study area where direct comparisons can be made. Abundance estimates with CVs in parenthesis.

	Visual survey			Digital video survey		
	Date flown	CDS estimates	DSM estimates	Date flown	Strip transect estimates	DSM estimates
Gulls						
WWT Stratum 1	20 March	1401 (11)	1347 (21)	18 March	1224 (64)	1173 (66)
WWT Stratum 4	18 March	1403 (20)	1691 (21)	22 April	103 (37)	100 (50)
WWT Stratum 6	18 March	1629 (22)	3521 (89)	22 April	542 (47)	441 (37)
Seabirds						
WWT Stratum 1	20 March	2998 (10)	2898 (09)	18 March	1257 (45)	1182 (45)
WWT Stratum 4	18 March	1020 (17)	981 (15)	22 April	21 (66)	12 (96)
WWT Stratum 6	18 March	3866 (20)	3870 (17)	22 April	804 (43)	660 (13)

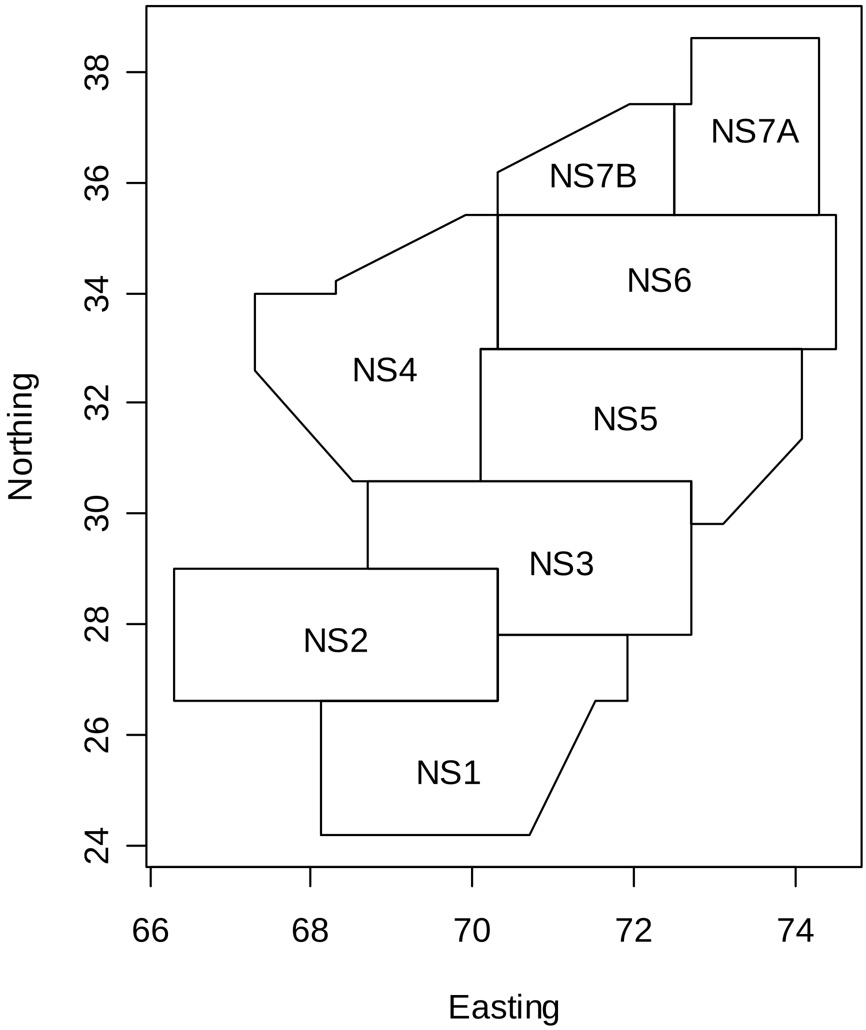

Figure 1. Norfolk round 5 survey region, broken into strata for visual surveys.

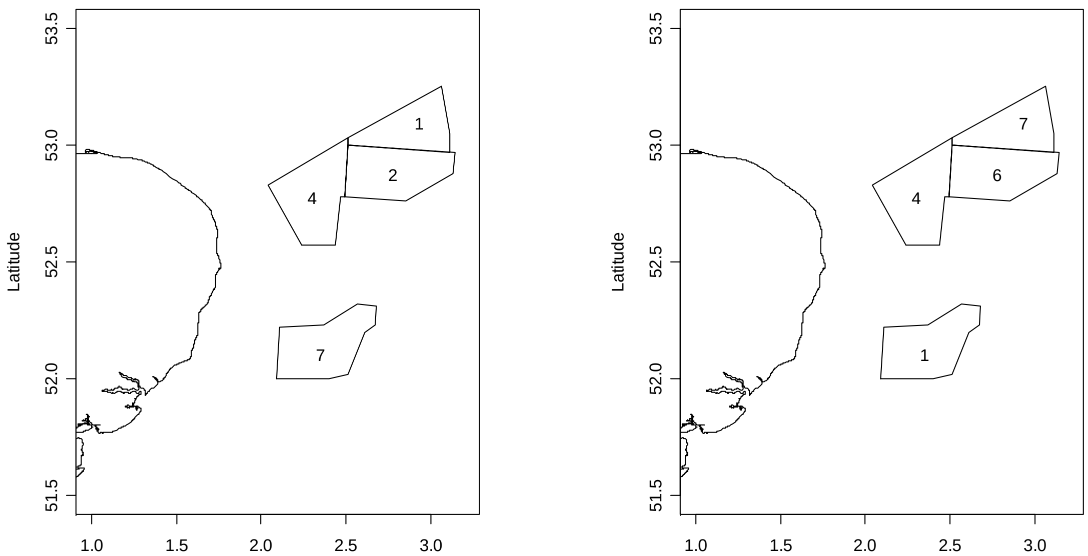

Figure 2. Survey block information provided for the digital video surveys; panel at left shows notation used as data were received; panel at right shows numbering to correspond with visual survey information.

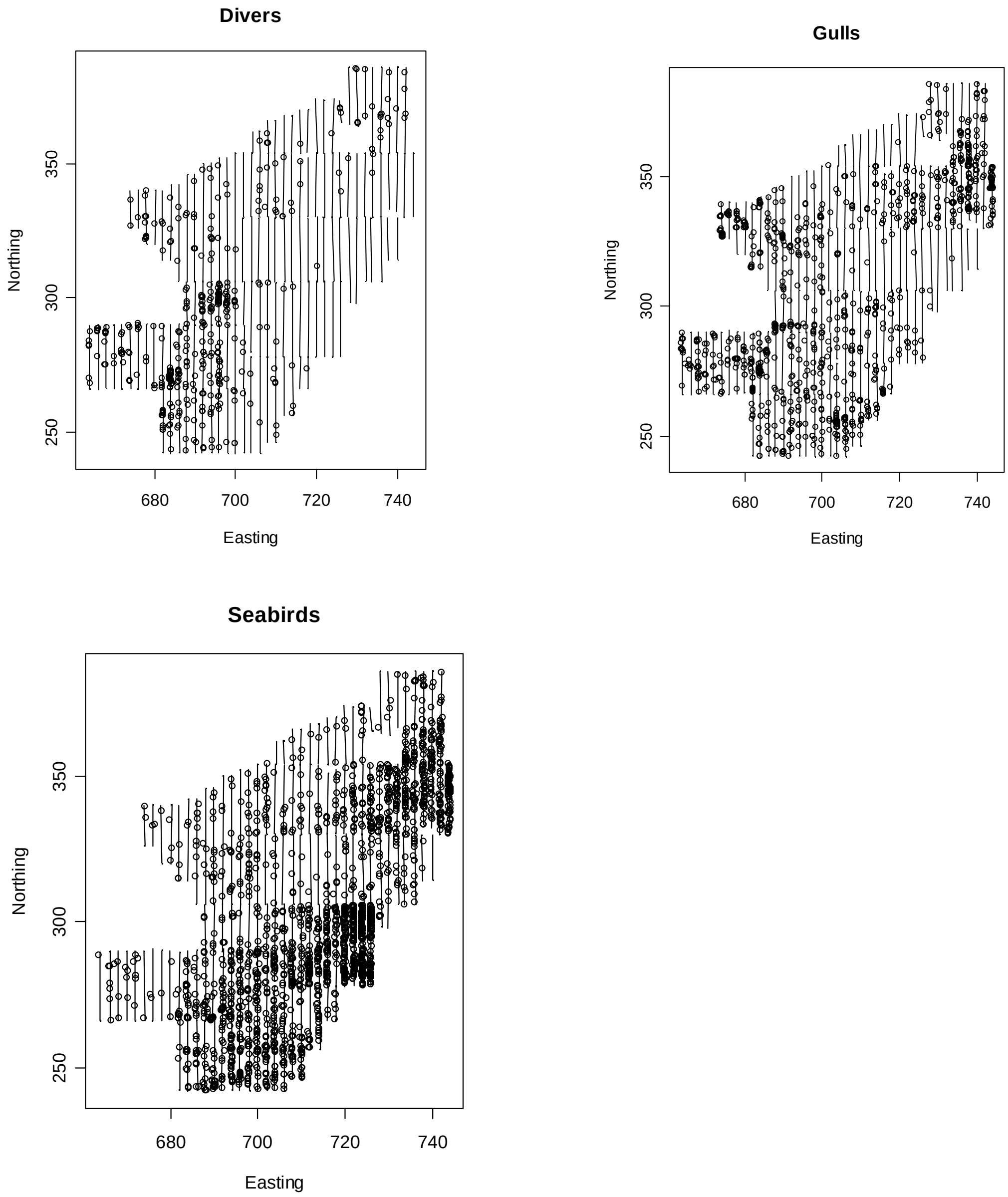

Figure 3. Map of survey effort (lines) and location of sightings for each species group (circles) derived from visual surveys.

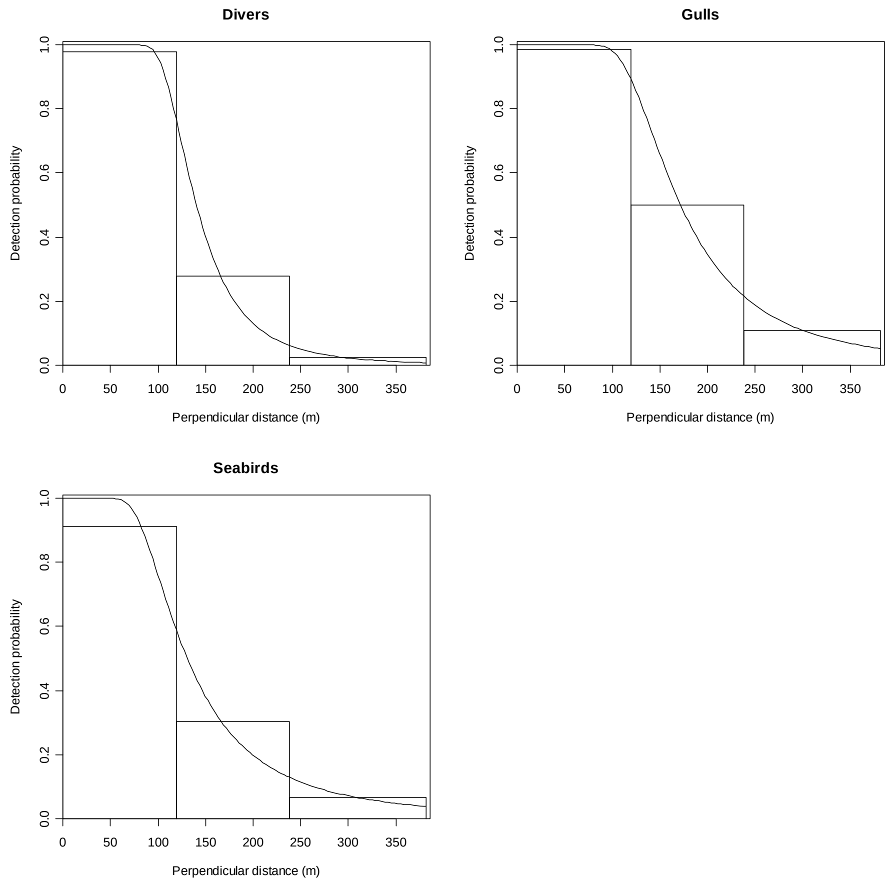

Figure 4. Estimated detection functions for each species group overlaid onto the scaled histograms of perpendicular distance for the visual survey data.

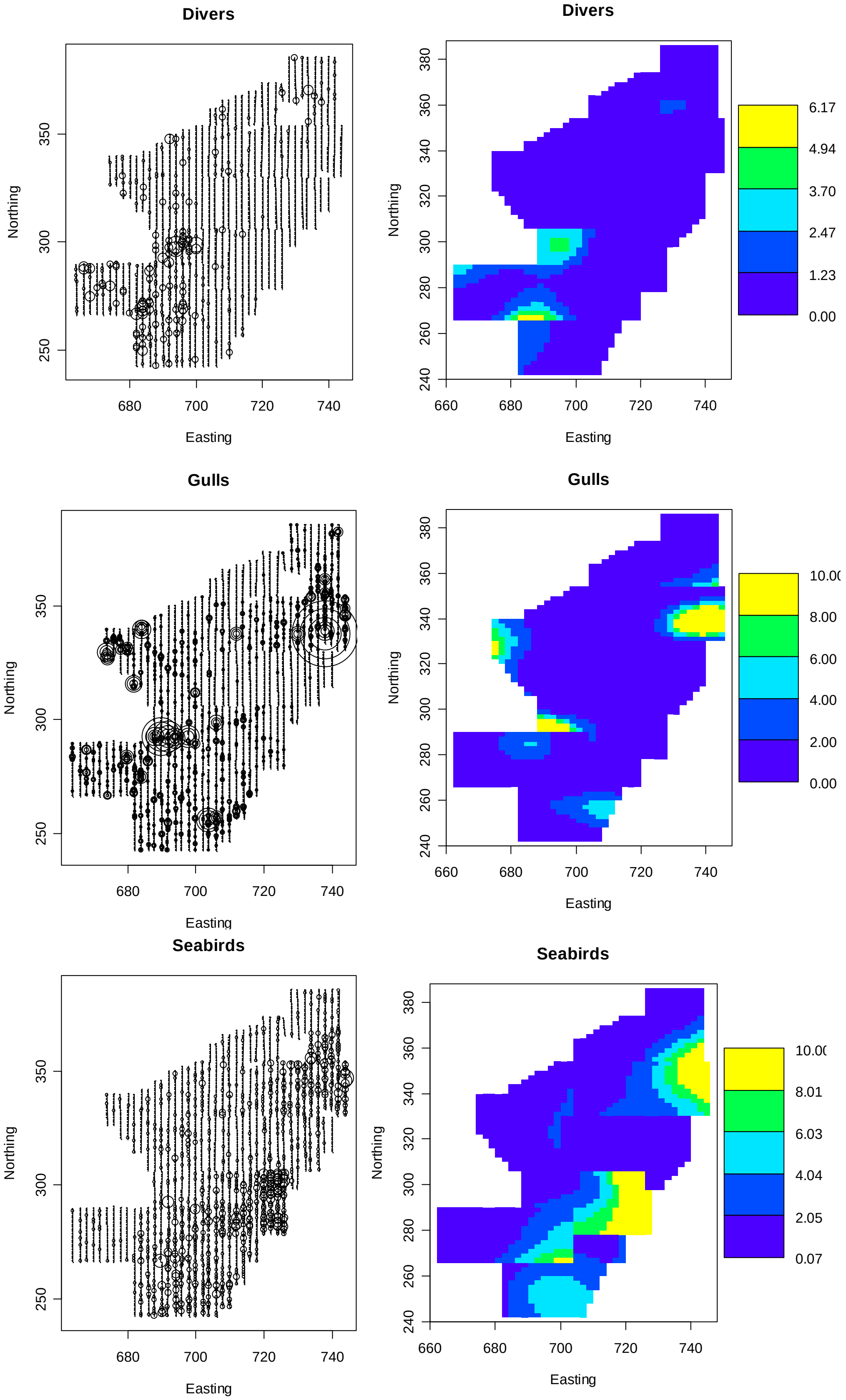

Figure 5. Estimated number of flocks per segment and fitted density surfaces (birds/km^{-2}) for each species group. The size of the circle is proportional to the estimated number of flocks; the scales are different for each species group. The scale bar for gulls and seabirds is capped at 10 birds·km^{-2} and so the last category includes higher densities.

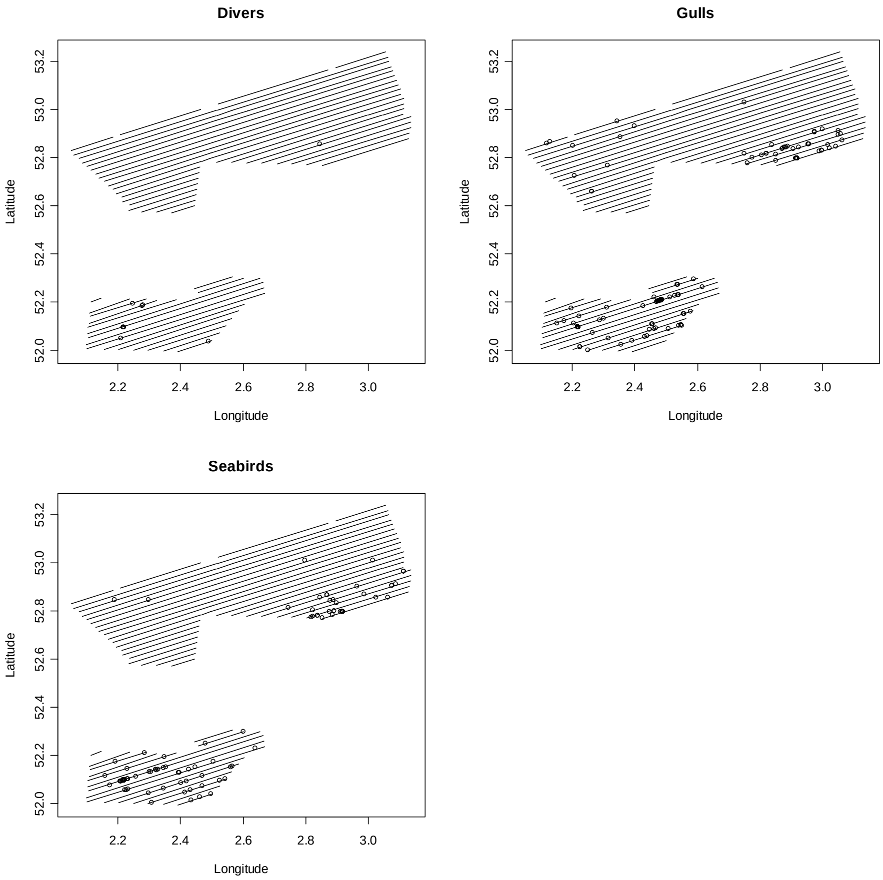

Figure 6. Transect lines and location of sightings (circles) by species group for digital video survey.

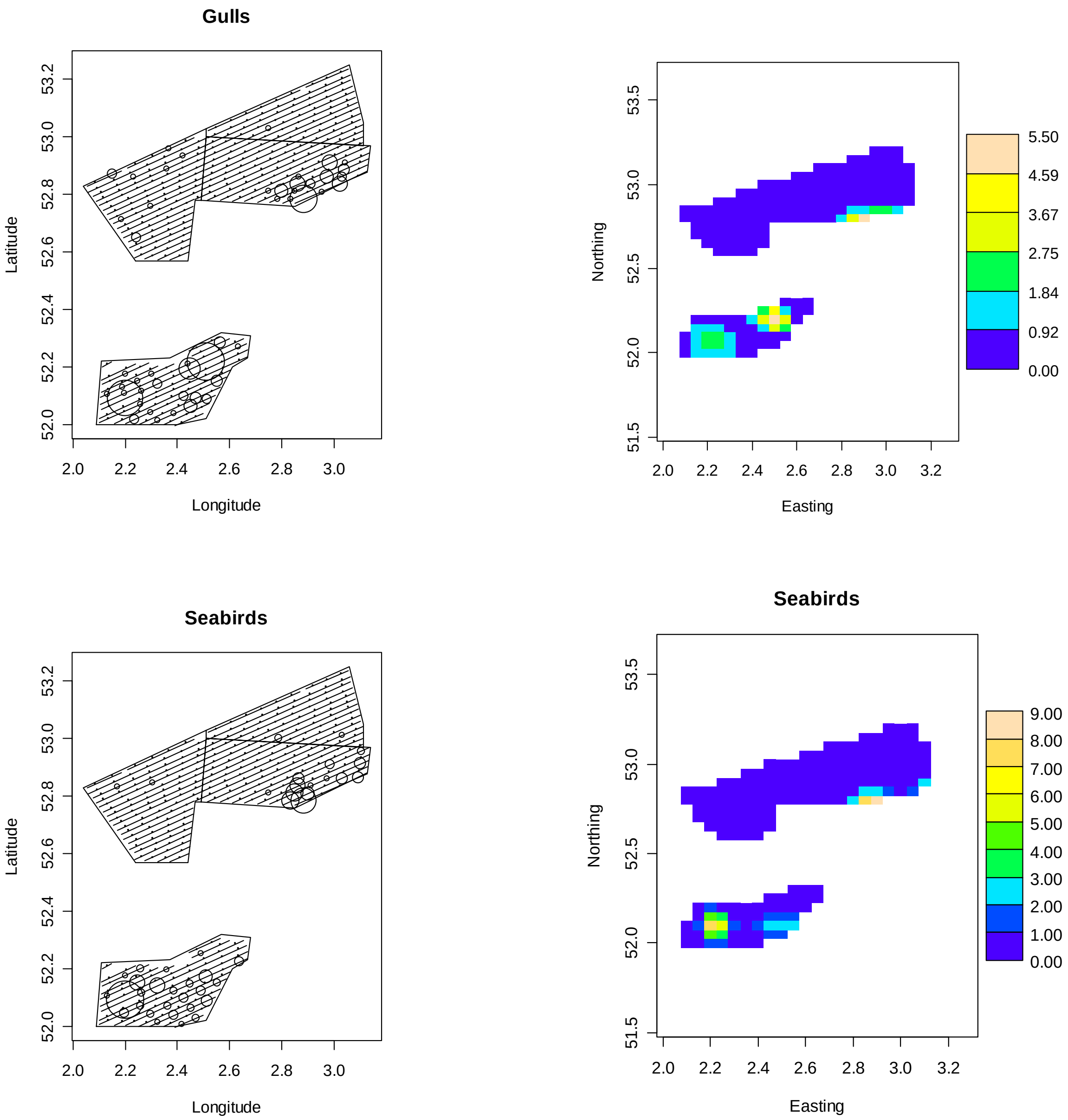

Figure 7. Density surfaces fitted to gulls and seabirds for the digital video survey data. Note too few diver detections were made to facilitate analysis of that species group for the digital video data.

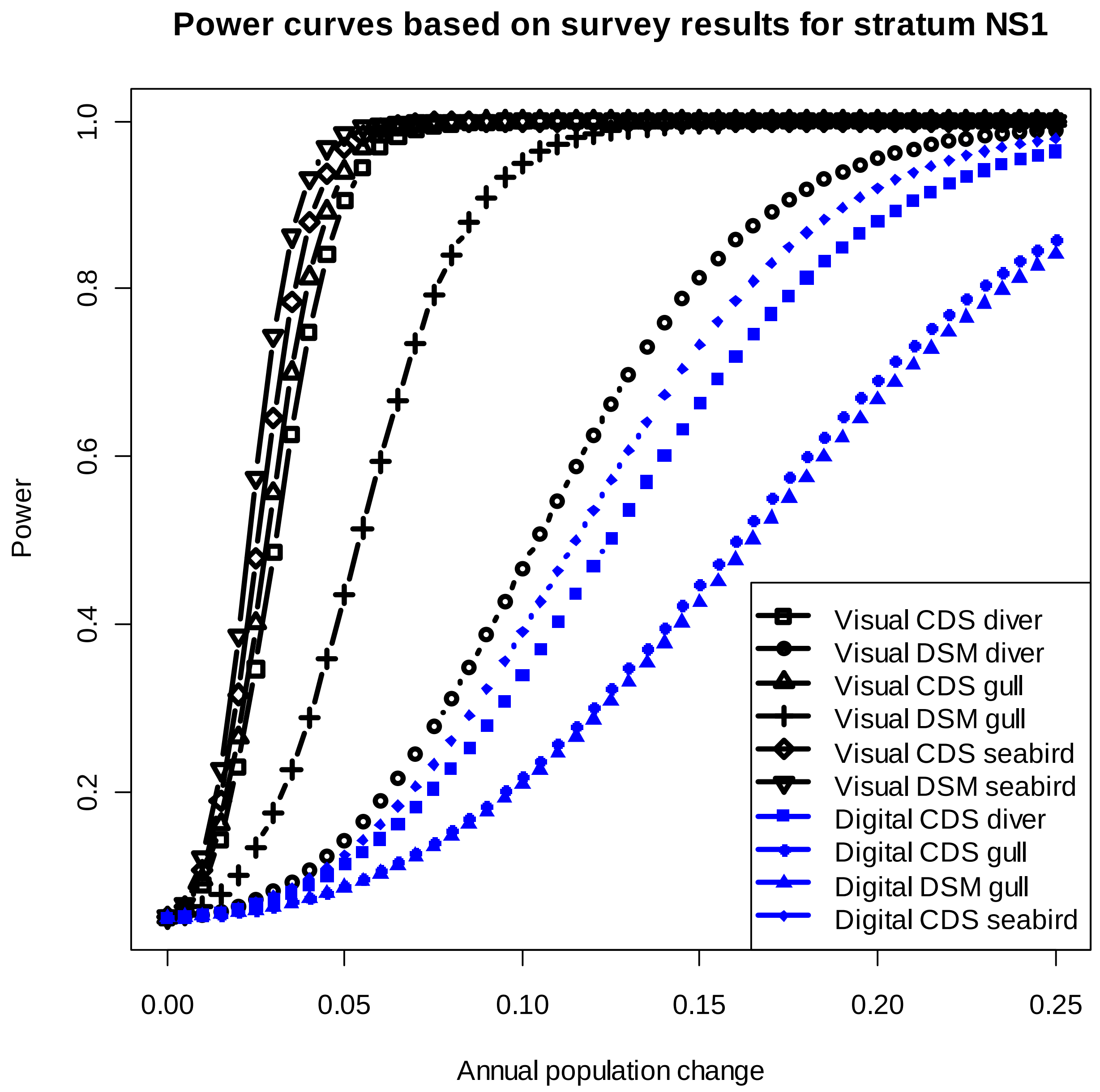

Figure 8. Power curves for survey/analysis/species combinations that were possible only for stratum NS1. Power for visual surveys in black, power for digital surveys in blue. Line style is the same for analysis type and species. CV (and hence power curves) for both analysis methods of the digital surveys for seabirds are identical (hence overlap for the highest of digital curves).

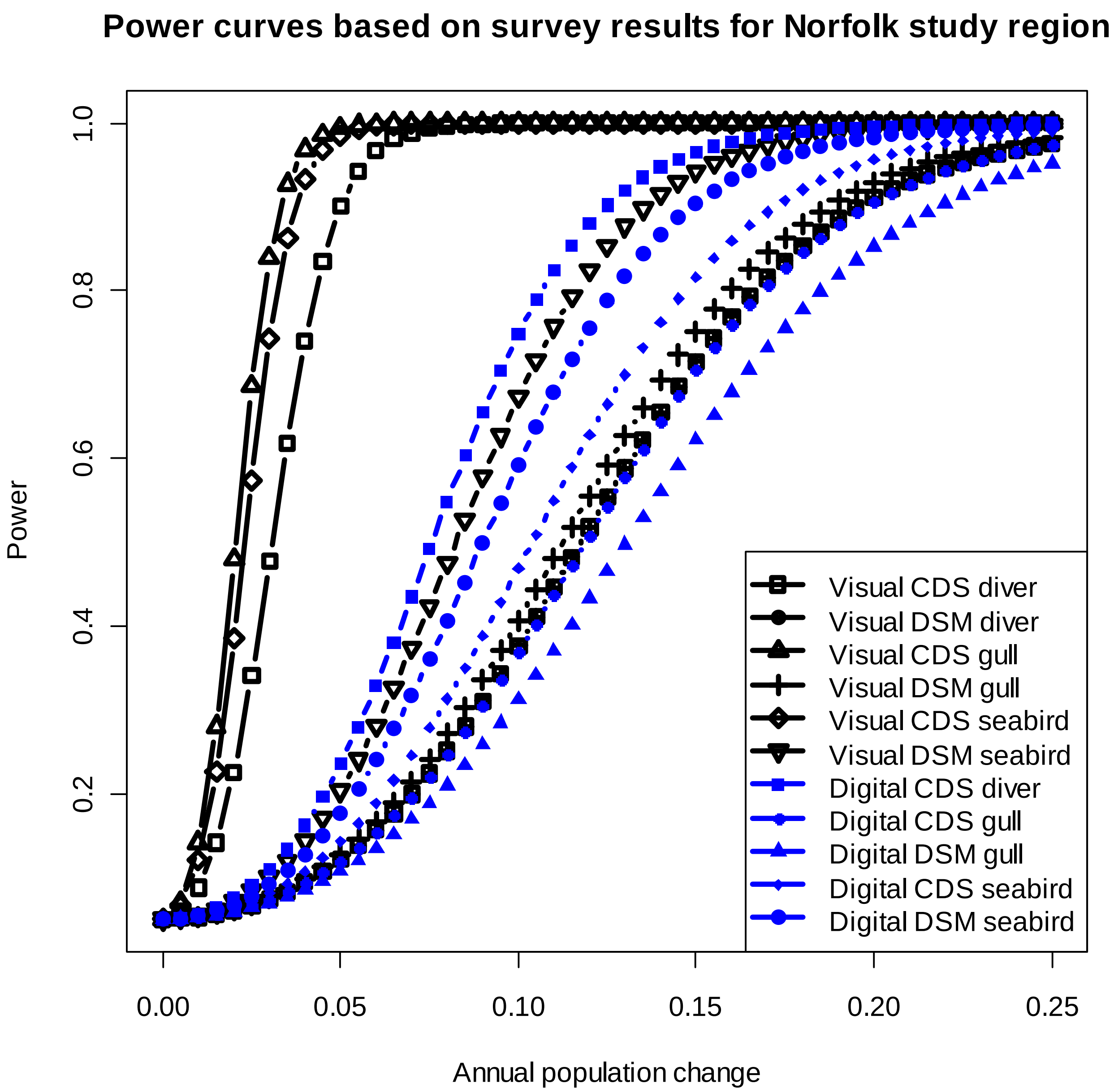

Figure 9. Power curves for survey/analysis/species combinations extrapolated to the entire Norfolk study region. Digital survey data for entire study region were not analyzed, so CVs were approximated by adjusting NS1 CVs by proportions equivalent to the reduction in CVs for visual surveys from the stratum to the entire study region. Power for visual surveys in black, power for digital surveys in blue.