COWRIE DISP-03-2006

BTO Research Report No. 446

Use of aerial surveys to detect bird displacement by offshore windfarms

Ilya M.D. Maclean[1], Henrik Skov[2], Mark M. Rehfisch[1] and Werner Piper[3]
[1]British Trust for Ornithology, The Nunnery, Thetford IP24 2PU, United Kingdom
[2]DHI Water and Environment, Agern Alle 5, DK-2970 Hørsholm, Denmark
[3]BIOLA, Gotenstr. 4, D–20097 Hamburg, Germany

October 2006

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

This report has been commissioned by COWRIE Ltd

Published by COWRIE Ltd.

ISBN-13: 978-0-9554279-0-9
ISBN-10: 0-9554279-0-8

Preferred way to cite this report:

Maclean, I.M.D., Skov, H., Rehfisch, M.M. and Piper, W. (2006) Use of aerial surveys to detect bird displacement by offshore windfarms. BTO Research Report No. 446 to COWRIE. BTO, Thetford.

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

Contact details:
Ilya M. D. Maclean
British Trust for Ornithology
The Nunnery
Thetford
Norfolk, IP24 2PU
Email: ilya.maclean@bto.org

Table of Contents

List of Tables and Figures

TABLES

Figures

Executive Summary

Offshore wind farms are likely to become one of Europe's most extensive technical interventions in marine habitats. European inshore coastal and offshore marine waters support globally significant numbers of seabirds and UK Government has legal obligations to monitor the effects coastal developments will have on populations of these species.

Aerial surveys potentially provide a cost-effective means of monitoring bird populations rapidly over large and inaccessible areas. However, the extent to which current survey protocols enable changes in bird numbers to be detected during wind farm construction and operation is poorly understood.

In this report we make use of existing aerial survey data and use power analyses to assess whether the current DTI aerial survey scheme can be used to assess whether changes in bird numbers occur, given that there are large background fluctuations in seabird numbers at any given site. Four taxa were selected for analysis: red-throated diver (*Gavia stellata*), common scoter (*Melanitta nigra*), sandwich tern (*Sterna sandvicensis*) and lesser and greater black-backed gulls (*Larus fuscus* and *L. marinus*). Aerial surveyors are not usually able to distinguish between these two large gull species. In addition, we tested the importance of using a higher resolution of collected distances in relation to the detection probability of target species during current DTI aerial surveys.

Increasing the number of distance bands used results in no perceptible reduction in the error associated with estimating detection functions using DISTANCE software. Greater precision is best achieved by increasing the number of transects flown over any given area, thus increasing the frequency with which birds are encountered.

Current aerial survey methods provide adequate means for detecting changes in the numbers for most species that are dispersed and not prone to large inter-annual fluctuations, like sandwich terns and black-backed gulls in the DTI data analysed. For those species, which are aggregated and prone to larger inter-annual fluctuations, like red-throated diver and common scoter in the DTI data analysed, existing aerial survey methods only provide restrained means of detecting changes in regions in which these species are particularly abundant.

Extending the duration of aerial surveys would increase the likelihood that changes in numbers could be detected, but not by a substantial amount. The probability of detecting change is influenced strongly by the average number of birds present and consequently a more efficient means of increasing the likelihood of detecting changes would be to increase the frequency of surveys at times of year when the target species are most abundant.

Analysing data using the same spatial-scale as that of the expected wind farm "footprint" and "buffer" maximises the probability of detecting changes in bird numbers. In order to distinguish between changes in bird numbers due to wind farm development from changes induced by other factors, changes within wind farm footprint and buffer areas should be compared to those in a nearby control or "reference" area (i.e. using a "before-after-control-impact" or BACI approach). Statistical comparison of changes between the footprint plus buffer and reference area increases the probability of detecting small wind farm induced changes within the footprint and buffer areas. However, the size of the reference area has little predictable effect on the likelihood of detecting changes in numbers. It is therefore advisable that selection of such reference areas is based on the biology and behaviour of the bird species present and not on the statistical likelihood of detecting changes.

Obtaining synoptic hydro-dynamic variables concurrently with bird data and incorporating these into analysis is likely to help explain some of the temporal variation

in numbers. Consequently doing so will increase the probability of distinguishing wind farm induced changes in bird numbers from background fluctuations. This method is likely to be the most cost-effective means of increasing the power of aerial surveys to detect changes in bird numbers.

Glossary

Accuracy - a term which refers to how closely an estimated value agrees with the correct value (a count of 55 birds is accurate but an estimate of 103 is not if in reality 56 birds are present).

Akaike's Information Criterion (AIC) - a statistic that assesses how well a statistical model fits, developed by Professor Akaike (Akaike 1976). Using a rigorous framework of information analysis, it also takes into account that a simpler model, i.e. one with fewer explanatory variables is generally better. In the context of this study it is used to determine which combination of covariates best explains bird count data and is calculated from the log-likelihood ratios obtained when undertaking generalized linear modelling, but also weights the statistic by the number of explanatory variables in the model.

Attraction – a term used to describe birds that are attracted to the footprint area and surrounding buffer area of a wind farm during its pre-construction, construction, operation or decommissioning, for reasons other than natural variability in bird population size and habitat quality. Such attraction could result from the presence of structures that are used for resting or from enhanced food supplies in the area of service vessels and turbines.

Covariate – a variable that is potentially predictive of the outcome under study, generally used in this report to refer to hydrological or environmental variables that are likely to have an effect on bird numbers and thus account for some of the variation in numbers. A static covariate is one that does not vary through time, whereas a dynamic covariate is one that changes through time, such as wind speed and water depth.

Degrees of Freedom - the number of independent pieces of information on which a parameter estimate is based and is a measure of the precision of the variance. The degrees of freedom for an estimate equals the number of observations (values) minus the number of additional parameters estimated for that calculation. As one has to estimate more parameters, the degrees of freedom available decreases. It can also be thought of as the number of observations (values) which are freely available to vary given the additional parameters estimated.

Detection probability function – function that describes the probability of detecting a bird or flock of birds at a given perpendicular distance from the transect line. In the software Distance, the detection function is modelled using a library of functions and series adjustments described in Buckland *et al*. (2001).

Displacement – a term used to describe the movement of birds from the footprint area and surrounding buffer area of a wind farm during its pre-construction, construction, operation or decommissioning, for reasons other than natural variability in bird population size and habitat quality. Such displacement may be caused by direct (e.g. loss of habitat) or indirect (e.g. habitat change, increased disturbance and noise due to maintenance activities) effects of the wind farm.

Explanatory variable - this variable (also called the independant variable) is the variable that is manipulated or selected by the experimenter to determine its relationship to an observed phenomenon (the response or dependent variable). In other words, a study will attempt to find evidence that the values of the explantory variables determine the values of the response variable (which is what is being measured). The explanatory variable can be changed as required, and its values do not represent a problem requiring explanation in an analysis, but are taken simply as given. In the context of this report, month, site or any of the covariates are explantory variables.

Generalized linear model - a statistical technique that allows one to calculate expected values from a set of observed values. In the context of this report it is used to estimate counts at any given site in any given year from observed data. The technique differs from ordinary regression techniques in that it allows the relationship between response variables and explanatory variables to be non-linear and can accommodate the response probability distribution being non-normally distributed as any member of an exponential family of distributions. Examples of the distributions that can be accommodated include negative binomial, Poisson, binomial and normal.

Logistic relationship - a mathematical relationship between the response variable and explanatory variables, which follows an S-shape, such that the initial relationship is exponential, but then slows and levels. It is also constrained such that the values of the response variable lie between zero and one.

Log-likelihood ratio - a statistical test relying on a computed test statistic, used in the context of this study to investigate how closely the observed counts in any given month at any given site compare to those predicted by a generalized linear model. It can be conveniently expressed using a simple formula including the Pearson's Chi-squared statistic.

Mean - is the average of a set of values obtained by adding them all together and dividing by the number of values.

Month effect - the influence a month has on bird numbers**.** It is estimated using the generalized linear modelling procedure and enables expected counts to be calculated for any given month (after accounting for site effects), using the formula: expected count = exp (month effect + site effect).

P-scale factor - when undertaking generalized linear modelling using a Poisson distribution (which assumes that the variance and mean are equal), it allows one to estimate the extent to which the variance may differ from the mean and thus the variance to mean ratio of a datasets (in context of this report, the inter-annual variance in count data in any given month at any given site), as the P-scale factor is the square-root of the variance to mean ratio. It is estimated from the ratio of the Pearson Chi-Square statistic to its degrees of freedom.

Parameter estimate - when undertaking statistical modelling, such as generalized linear modelling, the relationship between a response variable and one or more explanatory variables is sought. The parameter estimates are constants that give the extent to which the response varies as a result of changes in the explanatory variables.

Pearson Chi-Square statistic - is a value derived from one of the variety of statistical tests commonly used to evaluate how well observed values compare to predicted values. In the context of this report, it is a measure of how closely the observed counts in any given month at any given site compare to those predicted by a generalized linear model.

Precision – a term used to refer to the degree of confidence or known error range of an estimate (an estimate of 56 ± 1birds is precise, but an estimate of 50 birds ± 40 is less precise). An estimate can be more precise but less accurate than another (if 56 birds are present, an estimate of 92 ± 1 bird is more precise but less accurate than an estimate of 55 ± 5 birds).

Power analysis - an analytical technique used to determine statistical power. There are a number of ways in which it can be calculated, for example by rearranging the equation of a statistical test (see Cohen 1988 for a review of methods). In this study it is calculated by generating random datasets with the same characteristics (i.e. probability distribution, mean and variance) as real data, specifying a change in numbers (by adjusting the mean and variance), statistically analysing each dataset as if it were real data and then calculating the proportion of times that the statistical tests are significant.

Response probability distribution - the mathematical distribution of the response variable. Essentially when data are presented in a frequency histogram, if they approximate to a symmetrical bell-shaped curve they can often be assumed to be normally distributed. If the data are asymmetrical and positively skewed (i.e. more lower numbers, but high numbers differ from the mean by more than low numbers) then they can often be said to have a negative binomial distribution. A Poisson distribution is a special case of a negative binomial distribution in which the mean and variance are equal.

Response variable - this variable (also called the dependant variable) is the variable that is being measured and is affected by explanatory variables. In other words, a study will attempt to find evidence that the values of the explantory variables determine the values of the response variable. In the context of this report bird counts are considered to be be response variables.

Site effect - The influence a site has on bird numbers**.** It is estimated using the generalized linear modelling procedure and enables expected counts to be calculated for any given site (after accounting for month effects), using the formula: count = exp (month effect + site effect).

Statistical Power – or just power, is the probability of detecting a specified change in numbers. One minus the power (or beta) is the probability of falsely concluding that no decline has occurred when in fact a decline has occurred. In general results are expressed as a percentage that refers to the probability of detecting changes.

Statistical significance (alpha) – the probability of committing a 'Type 1' error, that is rejecting the null hypothesis (in this case of no changes in numbers) when it is in fact true.

Variance - a measure of the spread of the values in a group of numbers. The larger the variance, the larger the distance of the individual numbers from the group mean.

Variance to mean ratio - the variance divided by the mean. A frequency histogram of two sets of numbers with different means, but the same mean to variance ratio will have the same shape, only the one with the larger mean will be larger in size than the other. When performing statistical tests, to determine differences between two sets of numbers (e.g. bird numbers before and after the construction of a wind farm), statistical significance is affected by three things: (1) the sample size, (2) the amount of difference and (3) the variance to mean ratio.

Wind farm buffer area - an area surrounding the wind farm footprint area, in which the wind farm is thought to have an impact on birds.

Wind farm footprint area - the area in which a wind farm is situated.

Wind farm reference area - an area to which declines within the wind farm footprint and buffer areas are compared. Doing so, facilitates distinction between changes in bird abundance due to the wind farm development itself, rather than any potential confounding factors such as long-term changes in population size or distribution shifts in response to changing climate and weather.

Acronyms

AIC – Aikake’s Information Criterion
AVHRR - Advanced Very High Resolution Radiometer
BACI – Before-After-Control-Impact
BIOLA - Biologisch-landschaftsökologische Arbeitsmeinschaft
BTO – British Trust for Ornithology
CCW - Countryside Council for Wales
COWRIE – Collaborative Offshore Wind Research Into the Environment
DHI – DHI Water and Environment
DTI – Department of Trade and Industry
JNCC - Joint Nature Conservation Committee
NOAA– National Oceanic and Atmospheric Administration
OWF – Offshore Wind Farm
SNH – Scottish Natural Heritage
SST – Sea Surface Temperature

Units

Km – Kilometres
M – Metres
H - Hour
° - Degrees

1. Introduction

1.1 Background

Within the framework of the United Nations Climate Convention, industrial nations agreed in the 1997 Kyoto Protocol to reduce their greenhouse gas emissions by an average of 5% (compared to 1990) by 2012. The United Kingdom and other EU member states have committed themselves to reducing emissions by 8% (Exo *et al.* 2003). The EU White Paper on renewable energy aims at doubling the share of renewable energy by the year 2010, with a target of 40,000 MW from windpower. The UK government is committed to obtaining 10% and 20% of the UK's electricity from renewable sources by 2010 and 2020 respectively. Suitable land locations have become very limited and consequently major plans for offshore wind farms have been announced (Innogy 2003). According to current plans, within about 10 years, wind farms with a combined output of 40,000 megawatts will be installed in European seas, requiring an area of about 13,000 km^2 (Exo *et al.* 2003; Wind Directions 2003). Offshore wind farms are likely to become one of Europe's most extensive technical interventions in marine habitats (Merck and von Nordheim 2002; Exo *et al.* 2003).

Birds are likely to be one of the taxonomic groups most affected by wind farms through displacement from existing habitats or attraction to enhanced habitats (Exo *et al.* 2003; Garthe and Hüppop 2004; JNCC 2004; Desholm and Kahlert 2005). In the UK, all wild birds have a level of protection under the 1981 Wildlife and Countryside Act. Additionally, European inshore coastal and offshore marine waters support globally significant numbers of seabirds (Carter *et al.* 1993, Skov *et al*. 1995) and European Union Member States are obliged to protect populations of these species, under the EU Directive on the Conservation of Wild Birds (79/409/EEC, the Birds Directive) and the Ramsar Convention on Wetlands (Ramsar Convention Bureau 1988). These international agreements, together with the United Nations law of the Seas (United Nations 1982) and the EU Directive on the Assessment of the Effects of Certain Plans and Programmes on the Environment (2001/42/EC, the SEA Directive) require that states accept responsibility for assessing the effects of major offshore development on the environment. Governments are thus legally obliged to identify and designate the most important areas for birds as protected areas and undertake environmental impact assessments (EIAs) of the effects of developments in marine areas. Although ideally all species should be considered, it is generally recognised that assessment should concentrate on Schedule 1 species, Annex 1 species, regularly occurring relevant migratory species and species occurring at the site in regionally or nationally important numbers (SNH 2002).

Aggregations of large numbers of seabirds may be found in UK offshore waters throughout the year (Skov *et al.* 1995; JNCC 2004). These species differ in their sensitivity to offshore wind farms, and the extent to which they are likely to congregate at locations where offshore wind farms are proposed. The three development regions for offshore wind energy in the Thames Estuary, the Wash and the eastern Irish Sea span a relatively narrow range of oceanographical environments, as they are all shallow (< 30 m) areas with an almost mixed water column under strong tidal influence and in the case of the Wash and the Thames Estuary also under strong estuarine influence. Thus, the species of seabirds which primarily congregate in the three regions are estuarine species, which include divers, grebes, seaducks and terns. Pelagic species like northern gannet (*Morus bassanus*) and auks, which primarily congregate in more stratified and transparent waters, are also found regularly in all three regions. The most sensitive seabirds to wind farm developments in the three regions are red-throated diver (*Gavia stellata*) and common scoter (*Melanitta nigra*), both of which congregate in concentrations of relatively high importance and which due to their strong reactions to the presence of boats and offshore structures are easily displaced by disturbance (Garthe and Hüppop 2004; Mackey *et al.* 2000). Other sensitive species include eider (*Somateria mollissima*) and sandwich tern (*Sterna sandvicensis*) (Garthe and Hüppop 2004).

1.2 Monitoring methods

1.2.1 General

Typically, impact assessments for offshore wind farms follow three stages. Initially, a desk study and/or consultation exercise is undertaken to establish whether there are any bird populations at risk. If this exercise reveals that there are, then an evaluation of potential collision risk and direct and indirect disturbance for the relevant species is undertaken, and if necessary, population analysis to determine the likely impacts of the wind farm. If detrimental impacts on the population are likely to occur, then there is a requirement for undertaking habitat enhancement measures to outweigh any possible adverse effects (SNH 2002). The most frequent method adopted to assess the impacts of offshore wind-farms is the Before-After-Control-Impact (BACI) design (Stewart-Oaten *et al.* 1986), in which bird monitoring is carried out both prior to (baseline) and after wind farm construction, and the results compared to those from a control area situated within the same region, but some distance beyond the influence of the wind farm. Whilst this method has been criticised (Underwood 1994), it remains the preferred means of assessing the impact of offshore wind farms on birds in the UK (JNCC 2004).

Typically, data are collected for two or three years before construction in the seasons in which birds are most likely to be present in significant numbers (JNCC 2004). Whilst generally this entails monitoring in winter, when large congregations of birds occur offshore, monitoring is often also necessary during the summer if the site is situated close to seabird colonies and during the passage period if sufficient numbers of birds are thought to pass through the area during migration (SNH 2002; JNCC 2004; Kahlert *et al.* 2004) and also, for example, in Germany (Federal Maritime and Hydrographic Agency 2003).

Baseline surveys should be sufficient to give a confident assessment of the numbers of birds present throughout the year and are generally conducted for at least two full seasons to give some indication of natural variability in numbers and distributions (Kahlert *et al.* 2000; JNCC 2004; Kahlert *et al.* 2004). Surveys should include a 1-2 km buffer as well as at least one control area of at least half the size of the proposed wind farm area more than 1.5 km from the nearest proposed turbine (JNCC 2004), although some Danish studies have suggested that birds may be displaced by 2-4 km (Petersen *et al.* 2004, Petersen 2005). The principal methods used to assess habitat use by seabirds in relation to offshore wind farms are boat and aerial surveys (Innogy 2003); the former is here discussed briefly and the latter in more detail.

1.2.2 Boat surveys

Ship-based surveys provide a high level of accuracy in species identification and in the assessments of the age and behaviour of seabirds present in a study area. This may be important in understanding the natural variability in seabird distributions (Camphuysen *et al*. 2004). Typically boat surveys entail undertaking line transects and recording the target species numbers, behaviour and flight direction (Innogy 2003; Camphuysen *et al*. 2004). Distances of the species from the boat are also recorded to account for declining detectibility with distance (Buckland *et al.* 2001, Durinck *et al*. 1994). This method provides the most reliable counts for most species, and provides a means for identifying behavioural responses of individual species of seabirds to the wind farm and the wide range of associated human activities. However, some species, such as red-throated divers and common scoters can flush at considerable distance ahead of survey vessels making it necessary to perform surveys with at least 3 observers and applying continuous scanning by binoculars and distance-angle corrections (Durinck *et al.* 1993; Camphuysen *et al.* 2004). However, often three observers are not used and there are additional problems, as no guidelines exist with respect to how the density of birds in flight should be calculated (Banks *et al.* 2006). Ship-based methods with the application of naked-eye search for birds on the water have been used extensively for surveying seabirds in the United Kingdom (Kaiser 2002; Camphuysen *et al.* 2004; JNCC 2004; Banks *et al.* 2005, 2006).

1.3 Aerial survey methods

1.3.1 History

In Europe, the use of aerial surveys to monitor species populations was formerly limited by high financial costs, but developed rapidly during the 1960s, particularly in Denmark (e.g. Joensen 1968; 1973; 1974). Historically, surveys were carried out at a variety of altitudes, dependent upon the weather conditions, habitat and species and were much constrained by the difficulties of navigating without the aid of global positioning technology (Camphuysen *et al.* 2004). Aerial survey methods were further developed in the mid-1980s (Laursen *et al*. 1997) and subsequently extended to cover much of the coastal areas of the Baltic Sea for wintering seaducks (Durinck *et al*. 1994) and UK waters (Dean *et al*. 2003). The protocols originally developed in Denmark are now standard practice in Britain, where aerial surveys are often carried out concurrently with boat surveys to, for example, cover coastal areas that are difficult to access by boat (Camphuysen *et al.* 2004). German and Dutch surveys were initially mainly conducted as part of the International Waterbird Census (IWC), but are increasingly used for seabirds and marine mammals. The German protocols (Diederichs *et al.* 2002) are similar to those used in the UK, but differ from the Dutch methodology (Camphuysen *et al*. 2004). With the wide availability of GPS and increased need to conduct fine-scale surveys for wind farm impact assessment, aerial surveys have become a much used method for obtaining bird data over large areas throughout Europe (Camphuysen *et al.* 2004).

1.3.2 Protocols for monitoring birds by aerial survey

Several guidelines for best practise when monitoring birds by aerial survey to assess the impacts of wind farms have been proposed by COWRIE (Camphuysen *et al.* 2004) and are now widely adopted as standard protocols. It is generally recommended that a twin-engine (for safety) high-wing aircraft be used with good all-round visibility for observers. In the UK and in Denmark a line-transect methodology is recommended in which the aircraft flies along transects 2 km apart at 185 km h^{-1} at 80 m altitude. In Germany, it is recommended that transects are flown at least 5 km apart at 76 m altitude and as slow as possible (Federal Maritime and Hydrographic Agency 2003). Flying transects at some distance apart and at low altitude minimises the risk of double counting, but enables high-resolution data to be collected. Subdivision of survey bands is recommended to allow calculations of detection probabilities. Typically three-bands are used (44-163 m; 164-432 m and 433-1000 m corresponding to inclinations in degrees from horizon of 60-25°, 25-10° and 10-4°). This number of bands appears is thought to be the best compromise between obtaining accurate density data and the short period of time available for cognitively processing and recording information. However, in some instances more bands may be advantageous, but since the best means of recording distances appears to be with an inclinometer, band divisions should correspond to sensible divisions in degrees from the horizon rather than actual distance. Two trained observers should be used, one covering each side of the aircraft, with all observations recorded continuously on a Dictaphone. The time of each bird sighting is recorded, ideally to the nearest second. Locations are later determined by cross-referencing these with a GPS track that is obtained throughout the flight with locations and times recorded at least every 5 seconds. Flights are normally conducted from aircraft as helicopters cause greater disturbance to wildlife (Camphuysen *et al.* 2004).

It is thought that the most statistically efficient study design is a set of line transects running perpendicular to a major environmental axis. This is advisable as many seabirds assort themselves according to food availability and water depth for example. Consequently a set of transects running perpendicular to the coast would be most appropriate as lines running parallel to the coast may incur sampling bias due to birds concentrating in a thin band, perhaps underneath the plane or concentrated in one particular distance band. With Distance sampling, the sample unit is often a single transect and consequently around 20 transects may be needed (Laursen *et al.* 1997). It is recommended that at least four flights of the whole area be undertaken during the winter (mid-October to mid-March), with counts carried out across the whole period if possible. Where breeding birds are present, at least three flights should be undertaken between May and July/August, with counts ideally undertaken in late May, late June and mid-July to early-August. Additional surveys may be required for any other periods considered likely to be important (post-breeding, moulting or spring/autumn passage) (JNCC 2004).

1.3.3 Relative merits of aerial surveys versus other survey methods

There are several advantages and disadvantages associated with using aircraft as a platform for seabird surveys. The speed of the aircraft guarantees rapid, near-simultaneous coverage of large areas, to provide a snapshot of distribution and density. It is also often cheaper to survey large areas by air than by boat. The disadvantage of aerial survey is that due to short observation time there will be identification problems and reduced count accuracy in terms of numbers and behaviour. Ship-based surveys provide a higher level of accuracy in species identification and assessments of age and behaviour of seabirds at sea - this may be important in understanding the natural variability in seabird distributions (Camphuysen *et al*. 2004). As a consequence of the high speed, aerial surveys possess a higher risk of missing the birds, which are in the water column and not on the surface during the passage of the plane. With the flying height of aircraft, a good perspective over an extensive area is provided, with an extended detectibility gradient. The downside is that there is only limited scope for collecting additional information on *in situ* biological, hydrographic or other environmental parameters, although GPS registrations enable subsequent analysis of bird distributions in relation to such parameters obtained by other methods (Laursen 1989; Laursen *et al.* 1997; Petersen *et al.* 2004; Camphuysen *et al.* 2004). Aircrafts are suitable for surveying most seabird species without causing excessive disturbance or attraction, at least prior to the arrival of the aircraft. This is not the case for ships, which may disturb species like red-throated diver at considerable distance ahead of the vessel, necessitating the extensive use of binoculars and distance-angle corrections to permit detection and some compromise of the survey method (Durinck *et al.* 1993; Camphuysen *et al*. 2004). A further advantage is that aircraft can work in very shallow or intertidal areas that are completely or virtually inaccessible to ships of the recommended size.

In general, aeroplanes provide efficient coverage with variable degree of error of most species over large areas being developed for wind energy, while boats provide more accurate coverage of seabirds and their behavioural reactions within the smaller impacted area of each offshore wind farm. Thus both methods should be viewed as complementary, in-so-far as each fulfils different objectives and hence ideally should be conducted or at least considered (Camphuysen *et al.* 2004; JNCC 2004).

1.4 The effects of wind farms on birds

1.4.1 General

There are five main ways in which a wind farm may affect the habitats of seabirds: (1) direct habitat loss taken by the turbine bases, (2) indirect habitat loss through disturbance from areas in proximity to the turbines; such disturbance may occur as a consequence of wind farm construction work, or due to the presence of the wind farm close to nesting or feeding sites or in habitual flight routes, (3) positive effects introduced by the colonisation of foundations and scour protections by sessile organisms and flora, which eventually may increase fish diversity during the operation phase and hence may increase the availability of potential prey for seabirds within the wind farm (Leonhard and Pedersen, 2004), (4) positive effects due to the presence of structures that could be used as perching, roosting or resting platforms and (5) Positive effects such as a wind farm making a contribution to reducing the effects of climate change (Kahlert 2000; SNH 2002; Exo *et al.* 2003).

1.4.2 Avoidance

Many migrating seabirds deflect their flight path as they approach offshore wind farms (Kahlert *et al.* 2004; Desholm and Kahlert 2004; Christensen and Houninsen 2005), often at a distance of c. 400-500 m from the wind turbines (Christensen and Houninsen 2005). However, study of this phenomenon is geographically limited, mainly to Denmark where eiders constitute a large proportion of the birds present. In the UK, only a small number of specialist studies, such as the work on common scoter in Liverpool Bay (Kaiser 2002) have considered this question. In Denmark, by tracking the spatial migration pattern of waterbirds by radar it was found that the diurnal percentage of flocks entering the Rødsand wind farm area decreased by a factor of 4.5. At night, 13.8% of flocks entered the area of the initially operating turbines, but only 6.5% of those flew closer than 50 m to the turbines. During the day, these figures were 4.5% and 12.3% respectively. This means that only 0.9% of the night migrants and 0.6% of the day migrants flew close enough to the turbines to be at risk of collision. The proportion of flocks entering the wind farm decreased significantly between pre-construction and initial operation (Desholm and Kahlert 2005; Kahlert *et al.* 2004). Many radar tracks disappear within close range of a wind farm suggesting that birds landed on the water or modified their flight path such that a smaller cross-sectional area was exposed to the radar beam (Christensen and Hounisen 2005). Visual observations of common scoter showed that of 96 flocks, 76 landed on the sea, 52 at a distance of >500 m from the wind farm and 2 closer than 300 m (Christensen and Hounisen 2005).

Overall however, assessing true avoidance rates and thus actual collision rates is hard, as corpses often drop into the water and wash away and radar cannot track avoidance very close to wind farms. Doing so is important however as accurate assessments of avoidance rates is critical when estimating rotor-blade induced mortality to birds (Chamberlain *et al.* 2006). Fortunately, recent advances in thermal and vibration detection systems suggest that it may be possible to do so automatically (JNCC 2004, Desholm *et al.* 2005; Petersen *pers. comm.*).

1.4.3 Disturbance and displacement

Disturbance by operating wind turbines can exclude birds from suitable breeding, roosting and feeding habitats and can effectively amount to habitat loss (Drewitt and Langston 2006). Whereas direct loss of habitat due to the foundations of the turbines seems to be of no major concern for birds, as the proportion of occupied sea surface is generally rather small (< 5 % of wind farm site), numerous studies have shown that wind farms may indirectly affect a much larger area. In terrestrial habitats, numbers of roosting and/or feeding birds decreased around the turbines up to a radius of 800 m (depending on the species), with migrants, particularly larger species, the most affected (Percival 1999; Hoetker *et al.* 2004). Offshore, similar displacements often occur. For example studies in Denmark revealed that over-wintering and staging long-tailed ducks *Clangula hyemalis* and common scoter were displaced by wind farms during the construction and initial operation period, as aerial surveys revealed higher densities during the baseline period than during either the construction or operational phase (Kahlert *et al*. 2004). Such displacement can have adverse effects due to over-crowding in remaining areas, which through density-dependent competition for food resources may lead to increased mortality. In Liverpool Bay, the most important site in the UK for common scoter (Collier *et al.* 2005), displacement of scoters from areas around North Hoyle, Rhys Flats, Burbo Bank, Gwynt-y-More and Shell Flat wind farms, is predicted to increase mortality from 7.3% to 11.7% (Kaiser 2002).

However, responses to disturbance vary. On land, disturbance effects have not been found in any species at distances in excess of 800 m from turbines (Pederson and Poulsen 1991; Gill *et al.* 1996; Percival 1998; Hoetker *et al.* 2004). In many cases the actual disturbance distance has been found to be very much smaller than this distance and in some instances no disturbance effects have been found at all. Breeding birds have not been found to be affected at a distance of more than 300 m from a turbine (Percival 1999; SNH 2002). At sea, disturbances are often greater with disturbance effects evident at distances of up to 4 km (Petersen 2005).

Divers and scoters are particularly vulnerable, as exemplified by the fact that they will avoid ships by as much as a few kilometres and are thus at particular risk of disturbance during the construction phase, but also during routine maintenance (Winkelman 1992; Exo *et al*. 2003).

However, determining the effects of displacement on bird mortalities is difficult, as the degree to which birds are affected by displacement is dependent upon how close neighbouring areas are to carrying capacity (Fretwell and Lucas 1970). The standard approach for assessing wind farm impacts can only estimate displacement induced mortality using scenario based approaches and habitat modelling (Banks *et al.* 2005; 2006). Methods that attempt to calculate mortality directly are rarely attempted, but are complex, typically entailing the use of field observations coupled with individual-based modelling approaches, such as that method used to predict the change in over-wintering mortality rates in common scoter due to wind farm avoidance in North Wales and Liverpool Bay (Kaiser 2002).

1.5 Effectiveness of current monitoring protocols

Baseline monitoring usually occurs for two - three years prior to construction followed by monitoring during the construction and operation phase. Whilst it is recognition that monitoring needs to be undertaken for sufficiently long to detect changes in bird density and abundance, the length of time required to achieve this is rarely tested (Camphuysen *et al.* 2004; JNCC 2004). Seabird densities between years can be highly variable and there are also significant seasonal, diurnal and spatial variability in numbers (Kahlert *et al.* 2000; Camphuysen *et al.* 2003; JNCC 2004; Banks *et al.* 2005; 2006) and as such, the current length of time over which monitoring occurs may not be sufficient.

Where tested, results suggest that the power of current survey designs may not be adequate to detect population changes (Innogy 2003; Sims *et al.* 2006). For example, the protocol used to monitor the impacts of the North Hoyle wind farm would only be sufficient to detect changes in excess of 100% for fulmar (*Fulmarus glacialis*), kittiwake (*Risa tridactyla*), red-throated diver and shag and in excess of 50% for other species with a statistical probability of 95%. Relaxing the statistical probability to 80%, would still only enable changes in excess of 50% to be detected for fulmar, kittiwake, red-throated diver and shag (*Phalacrocorax aristotelis*) and changes in excess of 35% for other species. For relatively small changes in the region of 10% - 25% survey efforts would have to be quadrupled (Innogy 2003).

1.6 Aims of this report

The primary aim of this report is to perform power analysis to assess the extent to which the current DTI aerial survey scheme can be used to assess whether changes in bird numbers occur during the construction and operation of offshore wind farms. Making use of this information, ways in which the methods used for counting species by aerial surveys and the subsequent analysis can be optimised, will be discussed. "Power analyses" are used to determine how much survey effort is necessary to detect predetermined changes in the numbers of seabirds and waterbirds that may occur in the footprint area, surrounding buffer area and the reference area of a wind farm during and after its construction. Power analysis (see Cohen 1988 for a detailed explanation) makes use of existing empirical data to determine, while in the process of designing an experiment, how large a sample is needed to enable statistical judgments that are accurate and reliable. Performing power analysis and sample size estimation is an important aspect of experimental design. If sample size is too low, the results will lack the precision to provide reliable answers to the questions it is investigating. If sample size is too large, time and resources will be wasted, often for minimal gain. Power analysis is an essential tool that makes it possible to determine from existing baseline data the frequency and timing of aerial survey data necessary to detect changes in bird numbers of a certain size at a particular significance level. For example, it can be used to calculate how much survey effort is required to be able to detect a 10% change in bird numbers with a 5% significance level, or a 50% change in bird numbers with a 20% significance level. It can be used to identify the optimal duration and timing of the survey effort necessary to detect significant changes in birds numbers (Innogy 2003). In addition, we tested the importance of using a higher resolution of collected distances in relation to the detection probability of target species during current DTI aerial surveys. The results of this test make it possible to discuss the potential for optimising the survey design by increasing the number of distance bands used or by using real rather than grouped distances. Throughout, the scope of work is considered in the context that the ongoing monitoring for displacement may be delivered in respect of a single offshore windfarm site. However, existing data from several survey sites are used for some of the proposed work as such data help determine monthly and annual variability in bird counts.

2. Methods

2.1 Aerial survey data

We selected four taxa for analysis: red-throated diver, common scoter, black-backed gulls and sandwich terns. As the majority of divers are unidentified to species level during aerial surveys, we added a proportion of the unidentified divers to the counts of red-throated divers equivalent to the proportion of positively identified red-throated divers relative to other identified diver species. The vast majority of scoter species were recorded as common scoter, presumably because although this species can be difficult to distinguish from other scoter species from the air, the areas surveyed are known to contain only small numbers of velvet scoter (*Melanitta fusca*). Thus additional unidentified scoters were not added to common scoter counts, although in reality the numbers of these are such that doing so would have not made any appreciable difference. Although a number of tern species were not identified to species level, these were not added to the sandwich tern counts, firstly because the number of unidentified terns was also relatively small compared to the number of sandwich terns and secondly because most unidentified terns are likely to refer to either common terns (*Sterna hirundo*), arctic terns (*S. paradisaea*) or roseate terns (*S. dougallii*) as these three species are much harder to tell apart from each other than they are from sandwich terns, which have a diagnostic structure in flight (Blomdahl *et al.* 2003). Since distinctions between lesser and great black-backed gulls (*Larus fuscus* and *L. marinus*) are rarely made by aerial observers, all observations referring to either one of these species and those referring to identified black-backed gulls were pooled.

For an evidence-based assessment of whether existing aerial survey programmes are suitable to determine whether wind farms affect bird distributions, it was necessary to determine by how much bird numbers and distributions fluctuate, as any change in distribution will have to be identified against this background of fluctuating counts. To perform such analysis, existing aerial survey data were used. We opted to use raw counts rather than data corrected for distance bias using DISTANCE (Buckland *et al.* 2001) to avoid introducing another element of uncertainty into the analyses. Coverage by existing aerial surveys used in analyses is shown in Figure 2.1. Table 2.1 shows the temporal coverage of these data.

Table 2.1. Temporal coverage of aerial surveys. Values denote the number of times aerial surveys were conducted in each survey area and time period.

	Winter					Summer				
	2001/02	2002/03	2003/04	2004/05	2005/06	2002	2003	2004	2005	2006
G1A	0	0	0	4	3	0	0	0	0	0
G1B	0	0	0	3	3	0	0	0	0	0
GW2	0	0	0	4	3	0	0	0	0	0
GW3	0	0	0	4	3	0	0	0	3	1
GW4	0	0	0	3	3	0	0	0	3	1
GW5	0	0	0	4	3	0	0	0	3	1
GW6	0	0	0	4	3	0	0	0	0	1
NW1	0	0	0	3	3	0	0	0	0	1
NW3	0	0	0	3	4	0	0	0	3	0
NW4	0	0	0	3	1	0	0	0	3	1
NW5	0	0	0	4	4	0	0	0	4	0
N6A	0	0	0	3	5	0	0	0	2	0
N6B	0	0	0	4	5	0	0	0	0	0
TH1	0	0	0	5	4	0	0	0	3	0
TH2	0	0	0	5	4	0	0	0	0	0
TH3	0	0	0	4	4	0	0	0	1	0
TH4	0	0	0	4	0	0	0	0	1	0
TH5	0	0	0	4	0	0	0	0	3	0
TH6	0	0	0	0	4	0	0	0	1	0
TH7	0	0	0	0	4	0	0	0	0	0
XX1	1	0	0	0	0	0	0	0	0	0
XX2	0	0	1	0	0	0	0	0	0	0
XX3	0	0	2	0	0	0	0	0	0	0
XX4	0	1	2	0	0	1	1	0	0	0
XX5	0	1	1	0	0	1	1	0	0	0
YY1	0	1	1	0	0	0	0	0	0	0
YY2	0	1	1	0	0	0	0	0	0	0
YY3	0	0	2	0	0	1	0	0	0	0
ZZ1	0	0	0	1	0	0	0	0	0	0
ZZ2	0	0	0	3	0	0	0	0	0	0
ZZ3	0	0	0	2	0	0	0	0	0	0
ZZ4	0	0	0	0	0	0	0	2	0	0
ZZ5	0	0	0	0	0	0	0	1	0	0
CX1[1]	0	4	5	0	0	0	0	0	0	0

[1] only common scoter data were collected from this area.

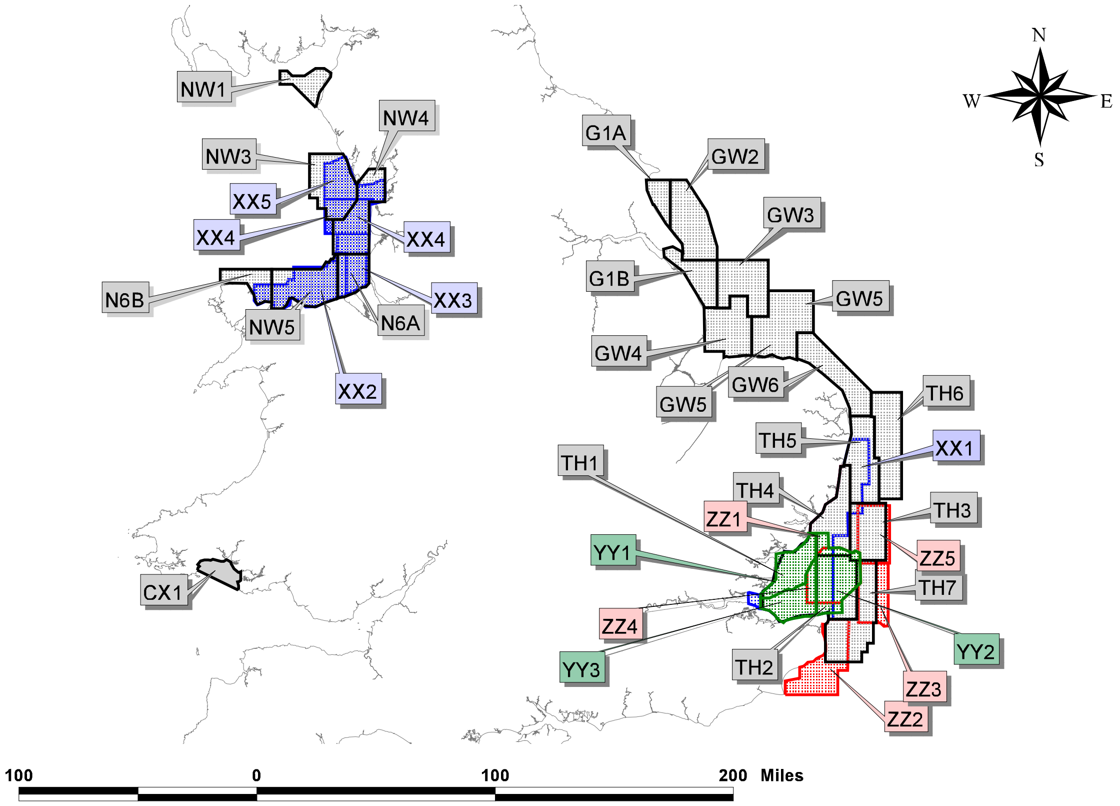


Figure 2.1. Survey areas covered by aerial surveys.

2.2 Power analysis

2.2.1 Overview

In the context of this study, statistical power is essentially the probability of being able to detect a specified change in numbers, in this case a change in bird numbers due to the construction of an offshore wind farm. Although there are a number of ways in which it can be calculated (see Innogy 2003 for an example relating to offshore wind farms and Cohen 1988 for a more general discussion), one common way of undertaking power analysis is to generate a series of random datasets with the same characteristics (i.e. mathematical distribution, mean and variance) as real data, specify a change in numbers (by adjusting the mean and variance), statistically analysing each dataset as if it were real data and then calculate the proportion of times that the statistical tests are significant (Cohen 1988).

In order to estimate the mean and variance of the real count data at any given site in any given month a generalized linear model was used. To fit the model, log-likelihood ratios (derived from Pearson's Chi-Square statistic), were used. The relationship between expected counts and explanatory variables was assumed to be logarithmic, which makes the assumption that predicted counts for any given month and site are related to the exponential of the site and month effects, thus constraining the predicted counts so that they cannot be negative. Count data were assumed to be Poisson distributed, but the degree of over-dispersion (i.e. the extent to which the count variances might be greater than the mean) was calculated from the ratio of the Pearson Chi-Square statistic to the degrees of freedom and is hereafter referred to as the P-scale factor. Both months and sites were considered to be categorical rather than continuous variables. The mean count for every site and month combination was calculated using the estimates for the site and month effects and the variance from the P-scale factor and mean as follows:

$$Variance = \frac{kq}{p^2}$$

where :

$$k = \frac{\mu p}{q},\ q = 1 - p \text{ and } p = \frac{1}{D^2}$$

where D is the P-scale factor and μ is the mean, given by exp (month effect + site effect).

By not performing separate analyses for each site, we make the assumption that the variance to mean ratio is constant across all sites and months. This was necessary, as we did not have a long-time series of data at our disposal for the majority of sites and thus calculating variance to mean ratios separately for counts at each site would limit the geographical scope of the study considerably and it would be unlikely that the results would be indicative of the full spectrum of sites. However for the few sites where more than 4 years of data were available, we tested this assumption by comparing actual variance to mean ratios to those predicted by assuming that they are constant across sites.

To simulate real data and thus perform power analysis, random counts for each site, month and year combination were generated. To represent data in years prior to construction, counts were generated using the site and month effects calculated from real data and by assuming a negative binomial distribution. To represent data in post-construction years, the mean and variance were adjusted by a specified range of declines (50%, 25% and 10%). The process was repeated ten times and statistical power calculated for each site by undertaking generalized linear modelling on each of the datasets and then calculating the mean proportion of times that specified declines could be detected as significantly different from zero, using a range of significance levels (0.05, 0.1, 0.2). Again, a logarithmic link-function and Poisson distribution with P-scaling was specified. Modelling using a Poisson distribution with P-scaling, accounts for the extent to which the variance may not be equal to the mean and thus allows a negative binomial distribution to be represented (Zar 1998). All computations were carried-out using SAS (SAS Institute, Cary, NC, USA). Although only a small number (10) of replicates were used to estimate the statistical power for each site, the statistical powers reported in this study are calculated as a mean across sites and given the large number of sites, estimates are based on a very large number of replicates, in excess of 10,000 using a scale of 10 km by 10 km for example.

2.2.2 Effects of spatial scale

When considering aerial survey data, it is rather arbitrary where one delineates the boundaries of a site and in theory data could be analysed at a range of spatial scales. To test the effects of spatial scale used for analysis on the statistical power of being able to detect wind farm displacement, statistical analysis was performed using a range of spatial scales. Initially, the following range of scales was used: 50 km x 50 km, 10 km x 10 km and 5 km x 5 km. Although the statistical power was greater when sites were larger for any given decline, there is generally a trade-off between the power of detecting displacement and the spatial scale at which data are analysed. Wind farm footprint and buffer areas are generally relatively small in size, sometimes as small as 4 km^2. Consequently, the decline due to a wind farm is likely to be relatively localised and is thus unlikely to be uniform throughout larger survey blocks as would be assumed in the power analyses if larger blocks are used. Consequently a 50% decline in a 10 km x 10 km site is not directly comparable to a 50% decline in a 50 km X 50 km site, but rather, is comparable to a 2% decline in the later (the ratio of the areas). Thus comparisons of power between sites of different scales, was performed by adjusting the decline as necessary (see Table 2.2). Subsequently, it was decided to use a spatial scale of 10 km x 10 km, as this scale gives a relatively high statistical power and is the same size as a typical wind farm footprint area and surrounding buffer. To determine the extent to which the spatial scale of analysis influences the variance to mean ratio, these were calculated for each species for the full range of spatial scales used.

Table 2.2. Spatial-scales and corresponding declines for which power analyses were performed

Scale	Area of site	Decline (percentage)
2 km x 2 km	4 km^2	100.00%
2.2 km x 2.2 km	4.84 km	82.64%
2.5 km x 2.5 km	6.25 km^2	64.00%
3 km x 3 km	9 km^2	44.44%
4 km x 4 km	16 km^2	25.00%
5 km x 5 km	25 km^2	16.00%
6 km x 6 km	36 km^2	11.11%
7 km x 7 km	49 km^2	8.16%
8 km x 8 km	64 km^2	6.25%
9 km x 9 km	81 km^2	4.94%
10 km x 10 km	100 km^2	4.00%

2.2.3 Effects of survey intensity

To test the relative merits of using just those survey months in which given species are abundant versus the value of using all survey months, results obtained by using winter months (January, February, March, October, November, December) only for red-throated diver and common scoter and summer months only (April, May, June, July, August, September) for sandwich tern are compared to those in which all months were used.

2.2.4 Effects of survey duration

Further power analyses investigated the effects of the length of time aerial surveys were conducted for 2 years pre-construction and 2 years post-construction, 3 years pre- and post construction, 4 years pre- and post-construction, 5 years pre- and post-construction and 10 years pre- and post-construction.

2.2.5 Effects of reference area

To circumvent issues associated with distinguishing changes in bird numbers due to wind farm displacement from changes due to larger-scale factors, declines or increases within a wind farm footprint (and surrounding buffer) area are normally compared to those within a nearby reference area. To determine the effect of undertaking such an analysis on the statistical power of detecting changes, analyses were performed using two types of reference area: one comprising four adjacent 10 km x 10 km survey blocks to the north, south, east and west of the "footprint and buffer" block and one comprising eight adjacent 10 km x 10 km survey blocks to the north, north-east, east, south-east, south, south-west, west and north-west.

As previously, random counts for each site, month and year combination were generated, 10 for each year using the mean and variance derived from real data for pre-construction years and scaled by the specified declines in the post construction years. Additional random counts with the same mean and variance of real reference area data in any given month at any given site. The methods used to calculate the mean and variance of real reference area data were the same as for other data. Power analysis was performed as previously, except that the proportion of times that the decline in the wind farm differed from the decline in the reference area was calculated. This was achieved by including an interaction term between a variable representing the effects of the wind farm footprint and buffer area and that representing the effects of the reference area.

2.2.6 Effects of mean and peak counts

To test the extent to which the statistical power of detecting changes is affected by the mean or peak number of birds present within a given 10 km x 10 km survey block, the relationships between mean power (α=0.05, 0.1 and 0.2) of detecting 50%, 25% and 10% declines within each block and the mean and peak counts logarithmically transformed were assessed using a generalized linear model, specifying a logistic relationship between logarithmically transformed mean and peak counts and a normal distribution. The logistic relationship constrains the power such that it could not be predicted to exceed 100% or be less than zero. Using the parameter estimates calculated from these models, the number of birds that would have to be present to detect such changes using statistical powers of 80% and 95% were calculated.

2.3 Effects of covariates

In general, one would expect that incorporating meteorological, topographic and hydro-dynamic covariates (i.e. ones which vary through time) into models would increase the likelihood of detecting changes as these are known to reduce the unexplained variability in the count data and thus the likelihood of being able to detect changes in bird numbers. However, we only had static covariates (i.e. ones which do not vary through time) at our disposal and consequently, by using the approach adopted in this study, whereby each site is considered on an individual basis, incorporating spatial covariates into the model that do not vary temporally cannot explain part of the variance in counts as such variance would be better explained by the incorporation of a site effect into the model directly. This is because every site would only be assigned one unique value of each static covariate and the variability in counts would be better explained by the characteristics of that site directly than indirectly through the use of static covariates. In order to improve the statistical power through the use of covariates, it would be necessary to incorporate hydro-dynamic covariates. As the incorporation of dynamic covariates was not possible within the time-frame of this study, we chose to investigate the effects these might have by comparing the reduction in the variance to mean ratio when static covariates are used to that obtained from models that do not include a site effect. We chose to do this rather than investigate the effects on statistical power directly as not including a site effect leaves so much unexplained variance in count data that the estimated power is extremely low. If the sample size and specified change in numbers does not differ between methods in which covariates are excluded or included, it is solely the variance to mean ratio that affects statistical power. Calculations of variance to mean ratios provides a better insight into improvements in power when expected power is very low and thus influenced heavily by the rounding errors that were necessary due to the low number of times that it was possible to repeat simulations within the specified time-frame of this study.

Six covariates were used: bathymetry (i.e. mean water depth), distance to land, distance to shallow water less than 10 m in depth, north slope aspect, east slope aspect and seabed complexity. These were obtained from DHI at a resolution of 10 km x 10 km and matched spatially to count data using ArcView (ESRI, Redlands, CA, USA). To determine which covariates were most appropriate to use for each species, models with all potential combinations of covariates were tested and that yielding the lowest AIC selected (Akaike 1976). Using the minimum AIC, red-throated diver counts were assumed to be best explained by the month, bathymetry, seabed complexity, distance from land and northern slope aspect. For common scoter counts were best explained the month, bathymetry, seabed complexity, eastward seabed aspect and northward slope aspect. For black-backed gulls the counts were best explained by the month, bathymetry, distance from land, northward slope aspect and distance from shallow water. For sandwich tern counts were best explained by the month, seabed complexity, northward slope aspect and distance from shallow water. The variance to mean ratio was then compared between the models that excluded covariate data (i.e. included a month effect only) and those that included the appropriate combination of covariate data.

2.4 Importance of resolution of collected distances

The current DTI aerial survey design, which has been adopted using techniques applied in Danish waters (Petersen *et al*. 2004), represents a line transect design which has been modified using the minimum number of perpendicular distance bands (3) required for estimating detection probability. The width of the three perpendicular distance bands has been selected on the basis of the most convenient inclinations from horizon (4°, 10°, 25°, 60°) equivalent to perpendicular distances of 44m, 164m, 433 and 1000m. Although exact recordings of perpendicular distances are only rarely possible during line transect surveys at sea, sampling of distances during most line transect surveys are such that the resolution of collected distances is considerably higher (Buckland *et al*. 2001). The collection and analysis of distance data with narrow intervals may improve the fit of the detection probability function, especially for data showing a tendency to spike near the transect line. In order to assess the potential optimisation of the current survey design by increasing the number of distance bands a comparative analysis of estimated detection probabilities was made using German aerial line transect data collected as real rather than grouped distances. The German data, collected in four different regions of the German Bight, allowed us to compare detection probability functions between the standard three band design and a design using perpendicular distance categories of 40 m.

Detection probabilities were estimated for 55 strata using the half-normal and uniform key functions with cosine adjustments. All possible combinations of adjustment function parameters were evaluated and the ones with the lowest AIC value were selected. Sufficient data for the analysis were available for the following species and species groups for which strata with more than 10 observations were selected: red-throated/black-throated diver, fulmar, black-backed gulls, kittiwake and auks. The results of the estimated detection probabilities for the two transect designs were evaluated by comparing the coefficient of variation of detection probabilities using a Wilcoxon matched pairs test.

3 Results

3.1 Power analysis

3.1.1. Modelled and actual variance to mean ratios

A comparison between actual variance to mean ratios and those predicted by assuming a constant ratio across sites, for the 15 10 km x 10 km areas for which data were available for more than 4 years is shown in Table 3.1.1. The ratio of the actual variance to mean ratio to that predicted (middle column) is an indication of the degree to which variability in counts are over or under-estimated by assuming variance to mean ratios are constant across all sites. The slope of the relationship between the actual variance to mean ratio and that predicted where both are logarithmically transformed (right-hand column) is an indication of the extent to which there is a general trend towards higher variance to mean ratios when mean counts are higher.

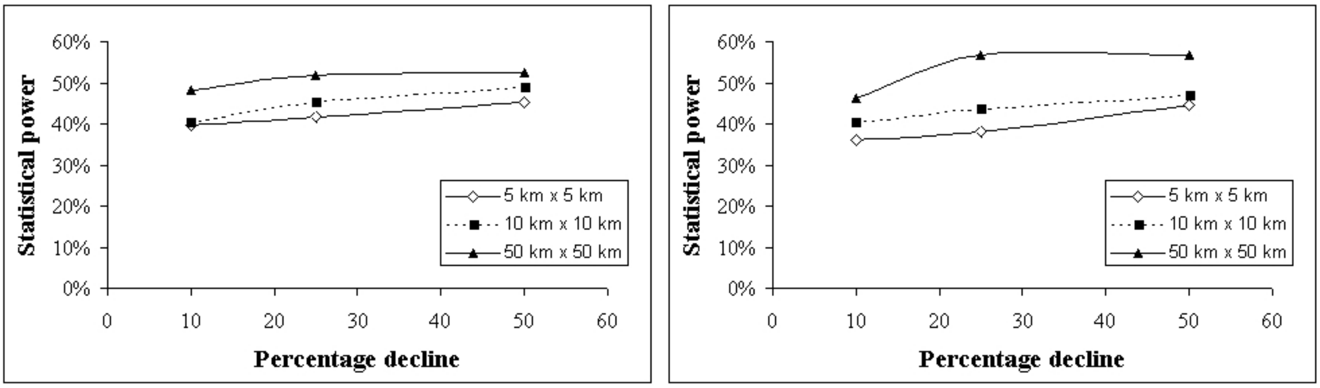
In general there was slight tendency for count variability to be over-estimated by assuming a constant variance to mean ratio across sites and that, with the exception of sandwich tern, actual site specific data reveal that the variance to mean ratio is generally higher when mean counts are greater. However it is uncertain as to whether these are real trends or merely artefacts arising due to the small number of potentially unrepresentative sites used to make this assessment.

Table 3.1.1. Relationship between actual variance to mean ratios of count data and those predicted by assuming a constant variance to mean ratio across all sites.

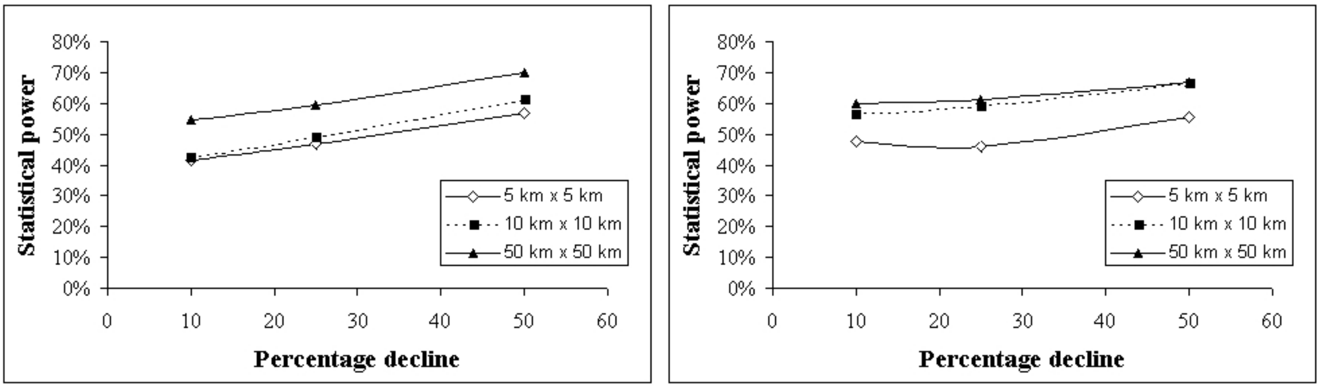
Species	Mean of actual variance to mean ratios / predicted variance to mean ratios	Slope of $\log_{10}$ actual variance to mean ratio = $\log_{10}$ predicted variance to mean ratio
red-throated diver	0.762	1.83
common scoter	0.764	1.90
black-backed gulls	0.746	1.59
sandwich tern	0.677	0.91

3.1.2 Effects of spatial-scale of analyses

The effects on statistical power of conducting surveys using different spatial scales are shown in Figure 3.1.1. In general, the larger the spatial scale, the higher the probability of detecting changes in bird numbers, although the actual differences in the statistical power obtained at scales from 5 to 50 km are moderate.



(a) red-throated diver (b) common scoter



(c) black-backed gulls (d) sandwich tern

Figure 3.1.1. Effects of the spatial-scale used for analysis, on the likelihood of detecting changes in bird numbers (statistical power), assuming data were collected monthly throughout the year for two years before and for two years after wind farm construction. Comparisons between the statistical power of undertaking surveys using 50 km x 50 km grids, 10 km x 10 km grids and 5 km x 5 km grids are shown for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern.

The decline due to a wind farm is likely to be relatively localised and unlikely to be uniform throughout larger survey blocks. Figure 3.1.2 shows comparisons of power between sites of different scales for a decline equivalent to 100% in a 2 km x 2 km grid cell (i.e. declines for cells were adjusted by the ratio of area their area to that of a 2 km x 2km cell: see Table 2.2). In general statistical power is greatest when small grid cells are used, but declines to a minimum when scales of c. 5 km x 5 km are used. For all subsequent analysis 10 km x 10 km grid cells are used.

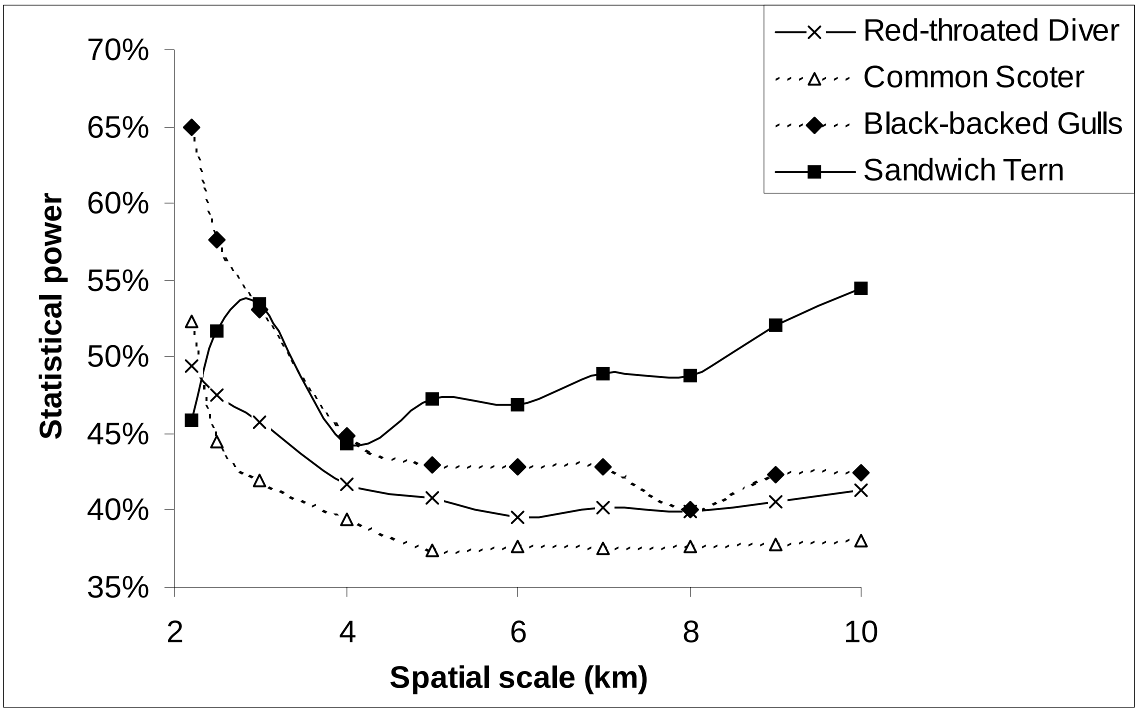


Figure 3.1.2. Effects of the spatial-scale used for analysis, on the likelihood of detecting changes in bird numbers (statistical power), assuming data were collected monthly throughout the year for two years before and for two years after wind farm construction. It is assumed that a decline equivalent to 100% within a 2km x 2km part of each survey block has occurred, but that no declines occurred elsewhere. Thus for any given scale, declines were calculated by dividing the area of the survey block at that scale by the area of a 2 km x 2km grid cell. Comparisons between the statistical power obtained using spatial-scales varying from 2.2 km x 2.2 km to 10 km x 10 km were undertaken.

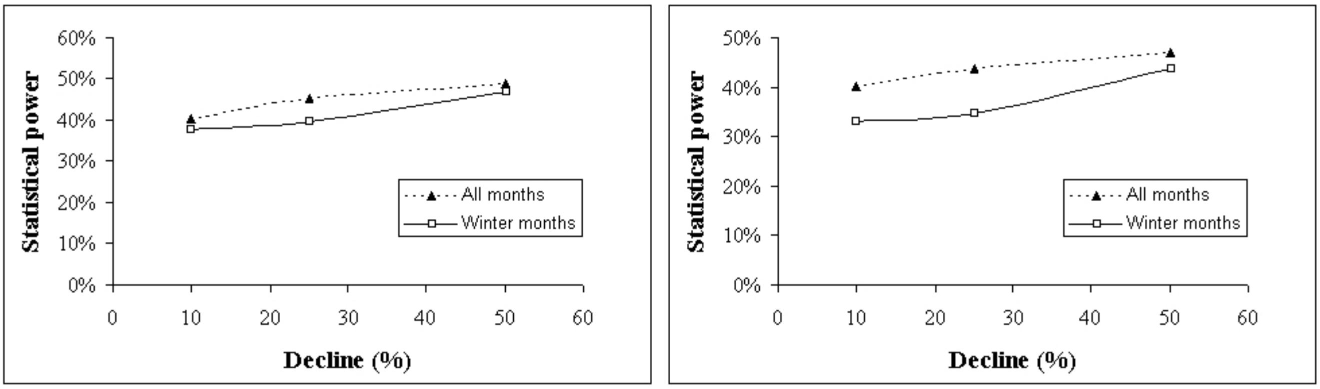
The extent to which spatial-scale affects the variance to mean ratio for each species across a range of spatial scales is shown in Table 3.1.2. This gives an indication of how variable the counts are (relative to the mean number recorded) at different scales and thus provides insight into why statistical powers are likely to be scale-dependent.

Table 3.1.2. The variance to mean ratio for each of the four species at a range of spatial scales (assuming a constant variance to mean ratio across sites and months).

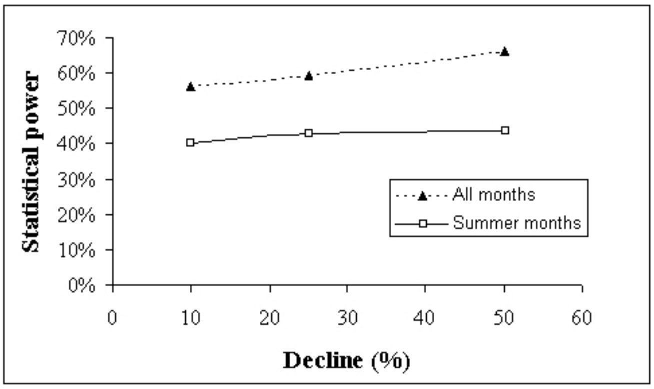
	red-throated diver	common scoter	black-backed gulls	sandwich tern
2 km x 2 km	27.06	1501.31	10.76	2.14
2.2 km x 2.2 km	34.91	1620.98	13.23	1.93
2.5 km x 2.5 km	37.31	1817.11	13.17	2.19
3 km x 3km	45.92	1811.92	15.09	2.42
4 km x 4km	59.79	2239.02	20.26	2.53
5 km x 5 km	79.10	2059.07	23.15	3.23
6 km x 6 km	84.58	2484.16	29.00	3.59
7 km x 7 km	115.91	2361.05	33.18	3.61
8 km x 8 km	122.42	2288.88	36.84	4.72
9 km x 9km	138.56	3080.35	41.64	5.21
10 km x 10 km	159.95	3937.05	50.85	6.08
50 km x 50 km	1289.27	6116.38	311.18	30.73

3.1.3. Effects of survey intensity

The effects on statistical power of conducting one survey every month in the year as opposed to just carrying out surveys for six months in winter (red-throated diver and common scoter) or summer (sandwich tern) are shown in Figures 3.3.1a-c. A 3 – 15% increase in the probability of detecting changes is evident when surveys are undertaken in all months. As black-backed gulls occur throughout the year, this analysis was not carried out for this taxa.



(a) red-throated diver (b) common scoter

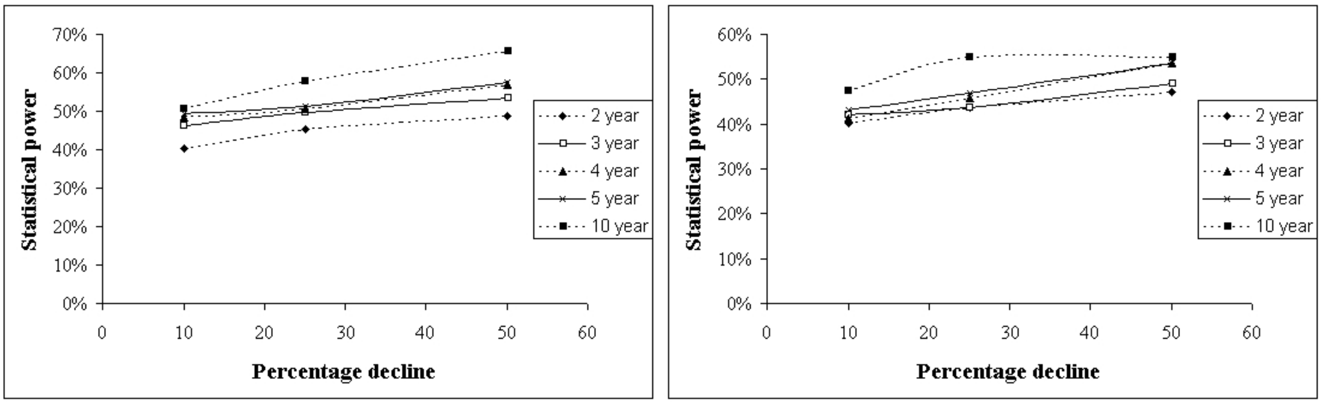


(c) sandwich tern

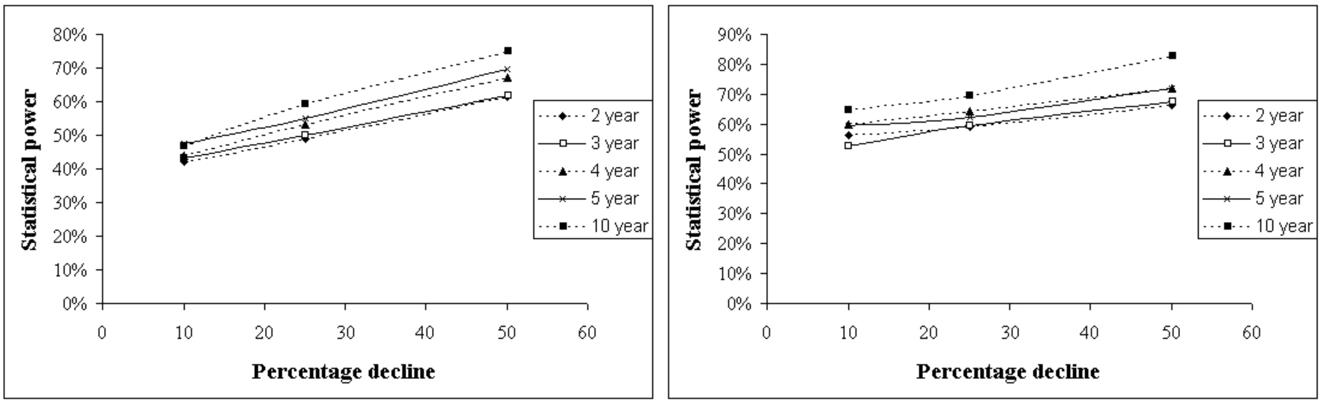
Figure 3.1.3. Effects of survey intensity on the likelihood of detecting changes in birds numbers (statistical power). Comparisons between the statistical power of undertaking surveys once every month throughout the year with the power of undertaking surveys in either winter months only for (a) red-throated diver and (b) common scoter or in summer months only (c) sandwich tern are shown. In both cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted for two years before and for two years after construction.

3.1.4 Effects of survey duration

The statistical power obtained by carrying out monthly surveys for different lengths of time are shown in Figure 3.1.4. In general, statistical power increased with survey duration, with the greatest improvement evident by extending surveys from two to three years pre- and post-construction for red-throated diver and from three to four years pre- and post-construction for other species.



(a) red-throated diver (b) common scoter

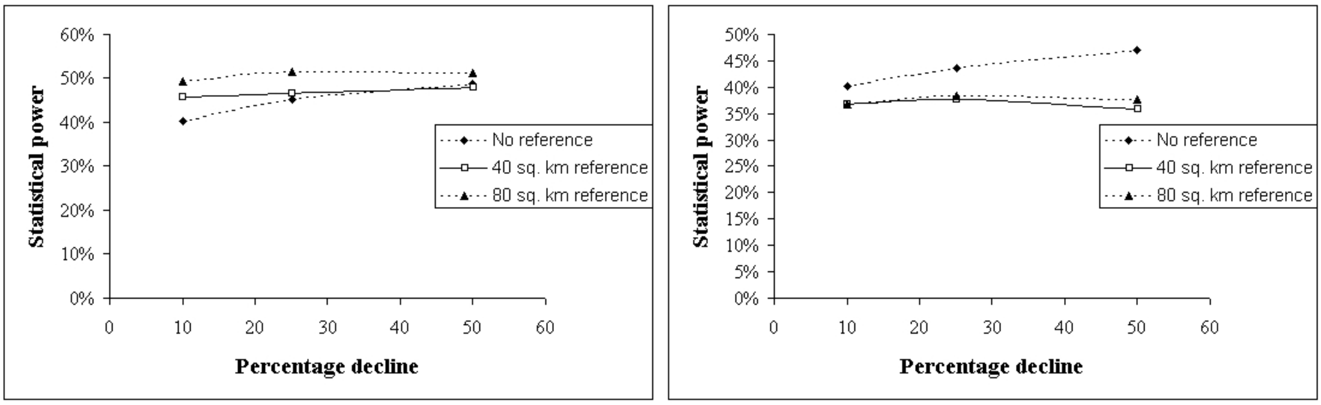


(c) black-backed gulls (d) sandwich tern

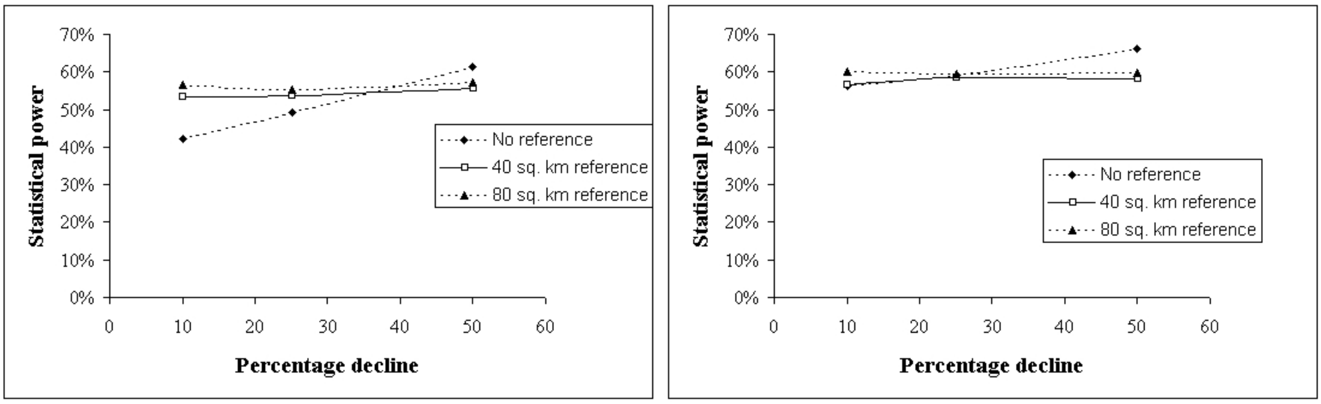
Figure 3.1.4. Effects of survey duration on the likelihood of being able to detect changes in bird numbers (statistical power). Comparisons between the statistical power of undertaking surveys from between two years and five years pre- and post-construction are shown for (a) red-throated diver and (b) common scoter or in summer months only (c) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year.

3.1.5 Effects of including a reference area

The effects on statistical power of including 40 km and 80 km control or reference areas compared to the effects when no reference area is included are shown in Figure 3.1.5. In general, the inclusion of reference areas increases the probability that small changes will be detected, but decreases the likelihood of detecting large changes. The increase in power with the inclusion of reference areas is, however, moderate, and is not seen for sandwich tern.



(a) red-throated diver (b) common scoter

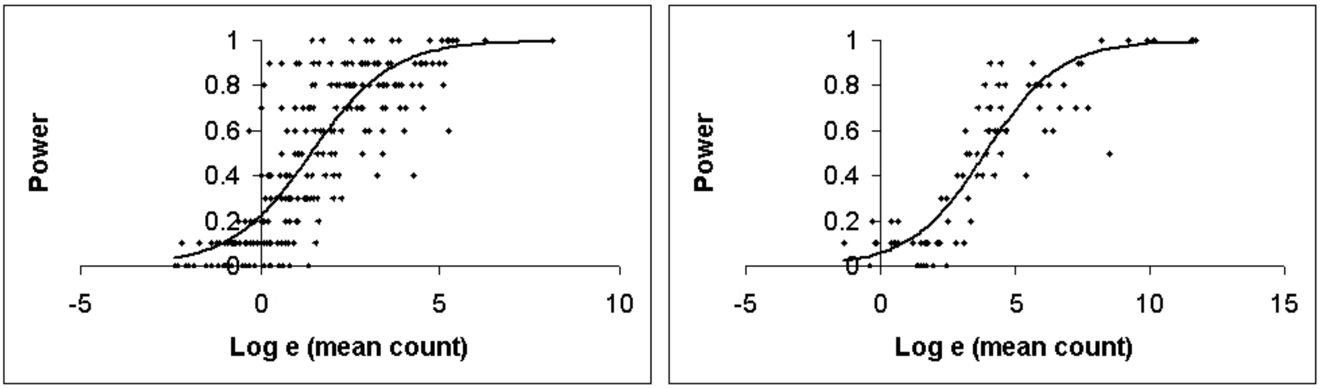


(c) black-backed gulls (d) sandwich tern

Figure 3.1.5. Effects of including a reference area on the likelihood of being able to detect changes in bird numbers (statistical power). Comparisons between the statistical power of including no reference area including four 10 km x 10 km cells and eight 10 km x 10 km cells are shown for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.

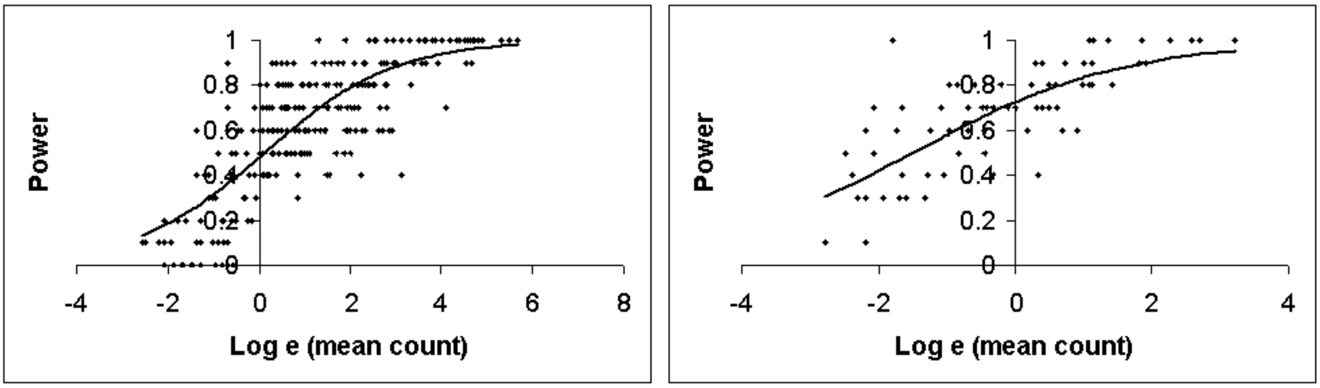
3.1.6 Effects of mean and peak counts

There was a strong relationship between the probability of detecting declines in birds and the mean number of birds on any given 10 km x 10 km grid cell. Figure 3.1.6 shows how the statistical power of detecting a 50% decline varies in relation to the mean number of birds present, assuming a significance level of 0.2. Figure 3.1.7 shows how the statistical power of detecting a 25% decline varies in relation to the peak number of birds present, assuming a significance level of 0.1. Figure 3.1.8 shows how the statistical power of detecting a 10% decline varies in relation to the mean number of birds present, assuming a significance level of 0.05. Figure 3.1.9 shows how the statistical power of detecting a 50% decline varies in relation to the peak number of birds present, assuming a significance level of 0.2. Figure 3.1.10 shows how the statistical power of detecting a 25% decline varies in relation to the peak number of birds present, assuming a significance level of 0.1. Figure 3.1.11 shows how the statistical power of detecting a 10% decline varies in relation to the peak number of birds present, assuming a significance level of 0.05. Table 3.1.2. shows the likely number of birds that would have to be present before specified declines in numbers could be detected.



(a) red-throated diver

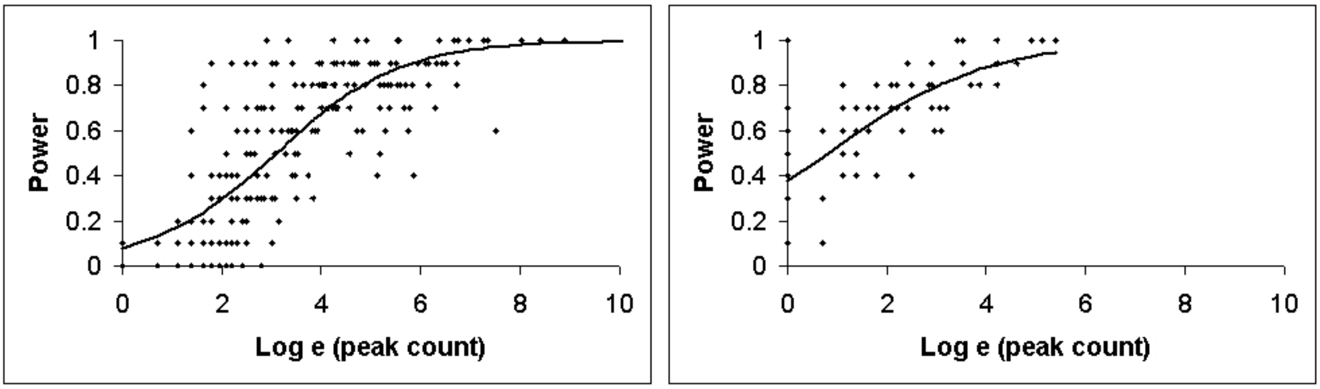
(b) common scoter



(c) black-backed gulls

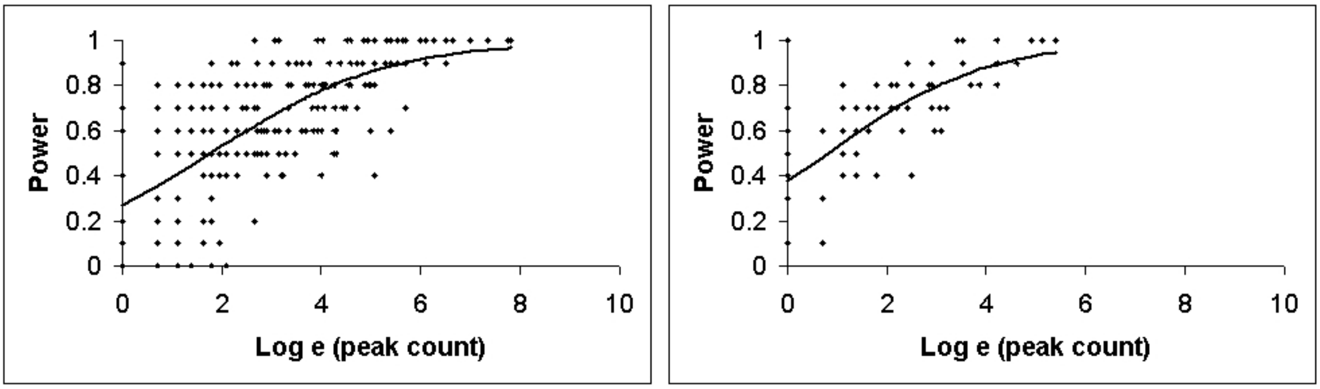
(d) sandwich tern

Figure 3.1.6. Relationship between mean number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 50% decline and statistical significance of 0.2, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.



(a) red-throated diver

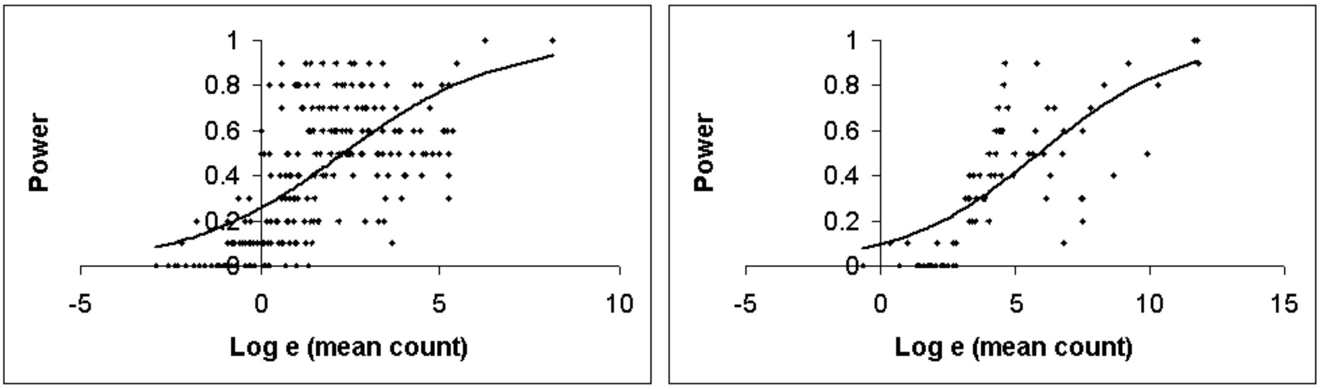
(b) common scoter



(c) black-backed gulls

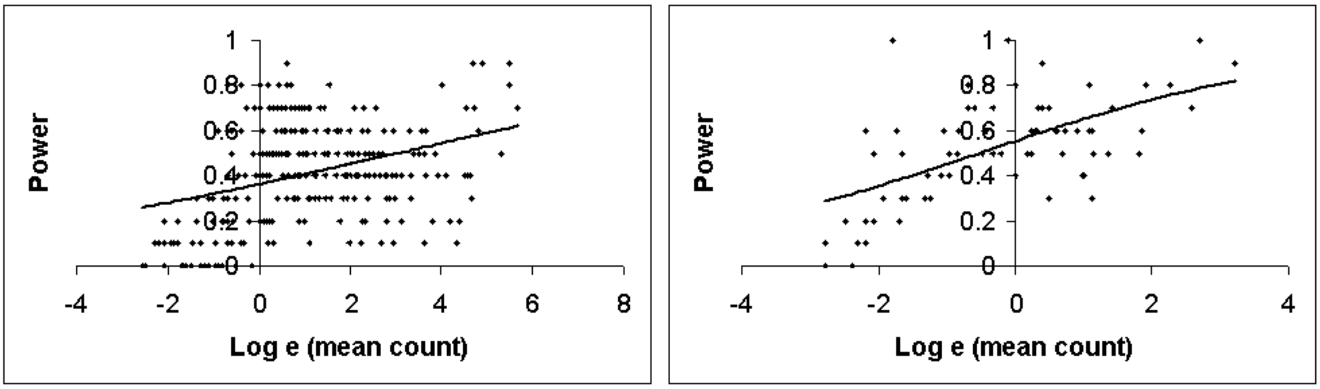
(d) sandwich tern

Figure 3.1.7. Relationship between peak number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 50% decline and statistical significance of 0.2, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.



(a) red-throated diver

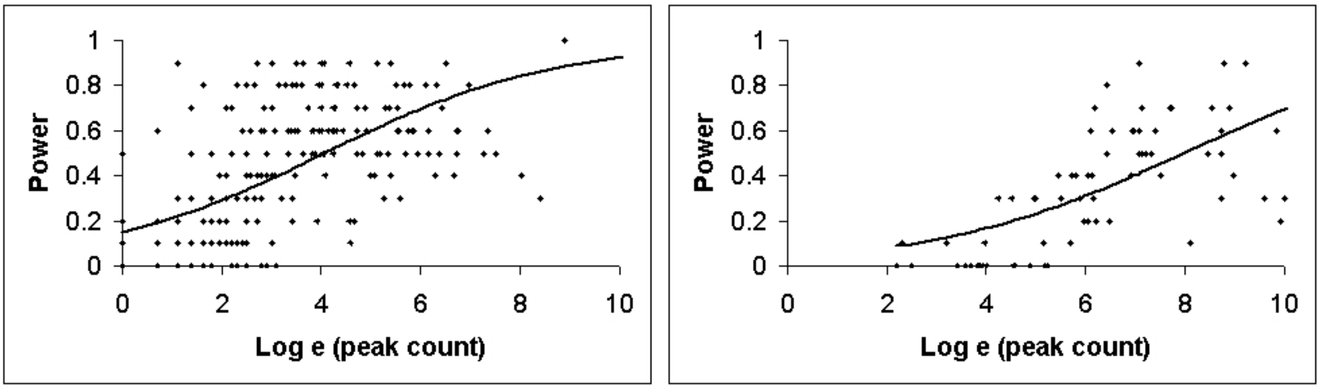
(b) common scoter



(c) black-backed gulls

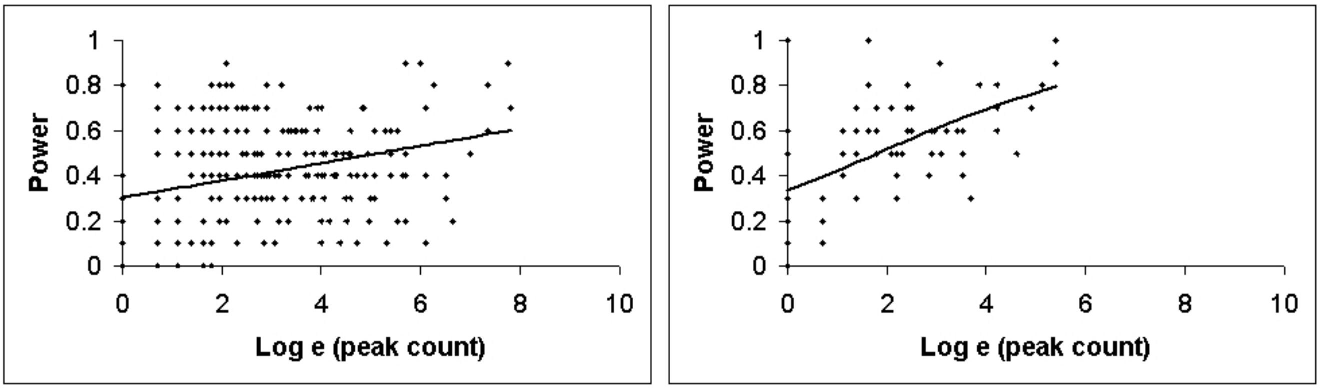
(d) sandwich tern

Figure 3.1.8. Relationship between mean number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 25% decline and statistical significance of 0.1, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.



(a) red-throated diver

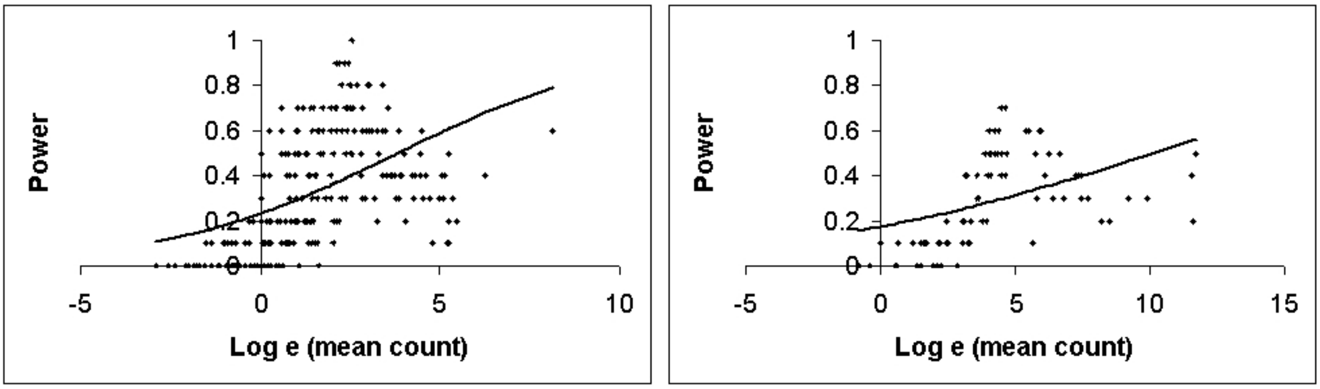
(b) common scoter



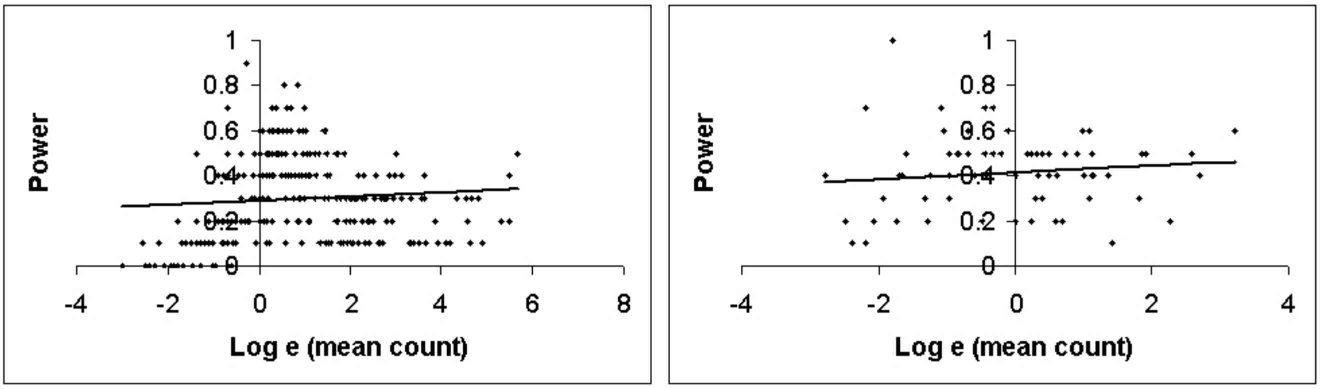
(c) black-backed gulls

(d) sandwich tern

Figure 3.1.9. Relationship between peak number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 25% decline and statistical significance of 0.1, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.

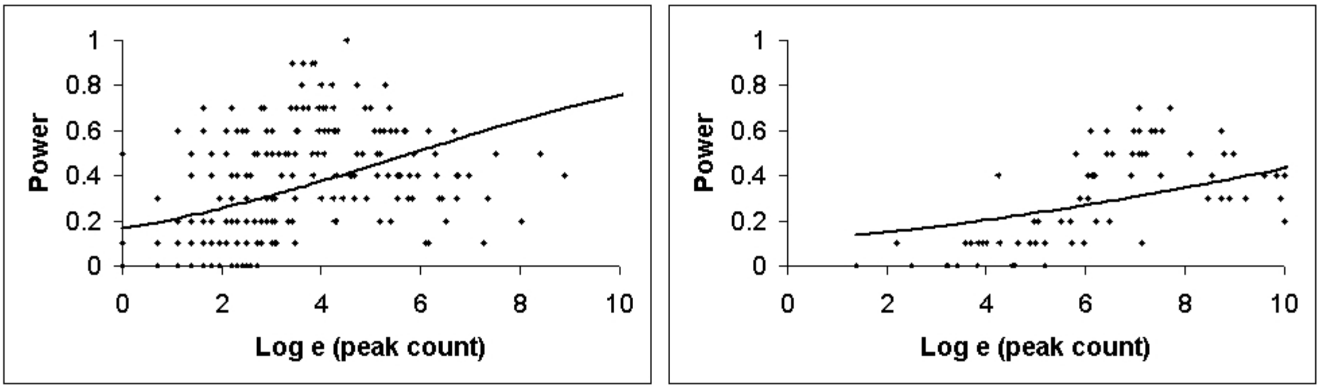


(a) red-throated diver (b) common scoter



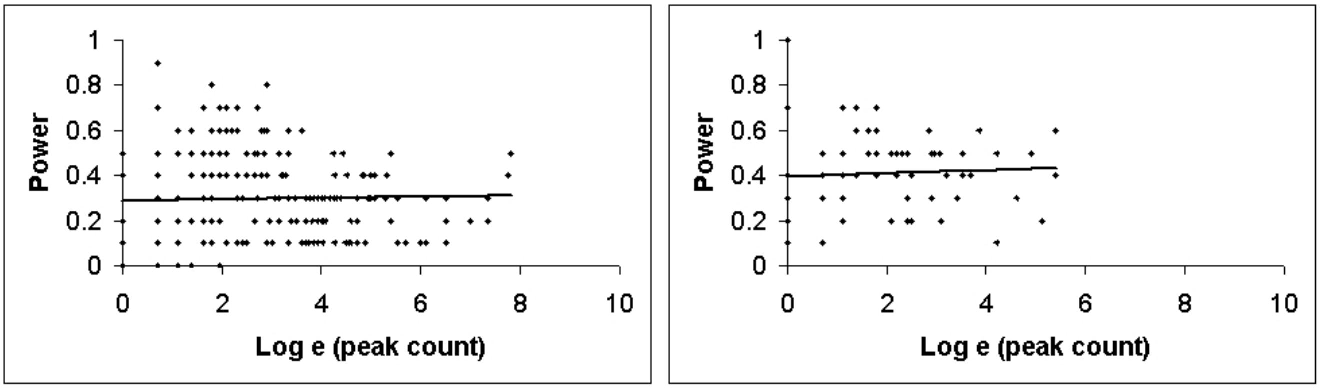
(c) black-backed gulls (d) sandwich tern

Figure 3.1.10. Relationship between mean number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 10% decline and statistical significance of 0.05, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.



(a) red-throated diver

(b) common scoter



(c) black-backed gulls

(d) sandwich tern

Figure 3.1.11. Relationship between peak number of birds present and the likelihood of detecting changes in bird numbers (statistical power) assuming a 10% decline and statistical significance of 0.05, for (a) red-throated diver, (b) common scoter, (c) black-backed gulls and (d) sandwich tern. In all cases data were analysed using a spatial resolution of 10 km x 10 km and it was assumed that surveys were conducted once a month throughout the year for two years prior to wind farm construction and two years after.

Table 3.1.3. Likely number of birds that would have to be present (pre-construction) before changes could be detected for specified declines. Alpha is the statistical significance – i.e. the probability of falsely assuming a decline when none is occurring. Beta is equivalent to one minus the statistical power and is the probability of falsely assuming no decline when in fact one is occurring.

	red-throated diver		**common scoter**		**black-backed gulls**		**sandwich tern**	
	β=0.2	**β=0.05**	**β=0.2**	**β=0.05**	**β=0.2**	**β=0.05**	**β=0.2**	**β=0.05**
50% decline, α=0.2 (mean count)	22	135	320	2,529	8	67	2	7
50% decline, α=0.2 (peak count)	126	893	4,434	43,725	66	266	20	85
25% decline, α=0.1 (mean count)	214	6,449	13,479	782,891	35,803	1.53×10^{8}	18	818
25% decline, α=0.1 (peak count)	1,501	57,491	88,982	4.17×10^{6}	1.22×10^{6}	2.46×10^{10}	241	15,021
10% decline, α=0.05 (mean count)	3,536	412,171	1.73×10^{8}	2.81×10^{12}	5.60×10^{22}	5.15×10^{37}	2.98×10^{11}	3.03×10^{21}
10% decline, α=0.05 (peak count)	45,670	9.82×10^{6}	1.54×10^{8}	5.19×10^{11}	5.37×10^{73}	2.17×10^{122}	7.18×10^{26}	2.01×10^{49}
National 1% threshold	49[1]		500		900[2]		105[3]	
International threshold	10,000		16,000		10,000[2]		1,700	

[1] 50 is normally used as a minimum threshold to define national importance

[2] Threshold based on summed thresholds for lesser and great black-backed gulls

[3] Threshold based on 1% of breeding population as given in Baker *et al.* (2005)

3.2. Covariate analysis

Although including static covariates do not improve the power of detecting changes in numbers over including site effects directly (see methods), models in which covariates are included have a substantially lower variance to mean ratio than those that do not. (Table 3.21). This would suggest that much of the spatial variation in bird counts derived from aerial data can be explained by oceanographical covariates.

Table 3.2.1 Comparison of variance to mean ratios between models that included static oceanographical covariates with those that did not.

		Scale factor	Variance to mean ratio	Ratio (without / with covariates)
red-throated diver	without covariates	50.82	3005.70	3.34
	with covariates	30.00	900.02	
common scoter	without covariates	82.06	6733.42	1.44
	with covariates	68.45	4684.80	
black-backed gulls	without covariates	17.05	290.70	1.24
	with covariates	15.29	233.78	
sandwich tern	without covariates	5.05	25.55	1.41
	with covariates	4.26	18.14	

3.3. Analysis of the importance of resolution in collected distances

The comparison between the standard transect design of aerial surveys with three distance band and a design using 40 m narrow bands revealed only minor improvements to the quality of the estimated detection probabilities with higher resolution of distance bands (Table 3.3.1).

In general, the CV of estimated detection probabilities only seemed to be lower for the high-resolution design when data displayed a clear spike in observations near the transect line. Most strata analysed did not exhibit a clear spike near the transect line, but rather displayed a wide shoulder with almost uniform detection probability to 100 m distance, followed by a sharp decline in detection probability between 100m and 200 m distance, and very low detection probability beyond 200m. Accordingly, no significant increases in the quality of the estimated detection probabilities were found by the introduction of the high-resolution transect design.

Table 3.3.1 Comparison of estimated detection probabilities between the standard aerial transect design and a high-resolution design with 40m distance categories. Analyses were carried out in Distance Ver. 5 and tested with Wilcoxon matched-pairs test on the basis of German aerial survey data. *P* indicates estimated detection probability within 400m.

			3 bands		10 bands	
Species	Stratum	Observations	P	CV %	P	CV %
red-throated/ black-throated diver	B10	12	0.239	24.9	0.216	21
	B8	25	0.276	16.7	0.239	14.5
	B9	40	0.272	13.2	0.262	11.4
	C6	35	0.435	34	0.307	22.6
	C7	24	0.298	16.9	0.26	14.7
	C8	61	0.303	8.8	0.295	20.4
	D11	30	0.314	15.2	0.372	25.3
	D12	46	0.288	7.4	0.251	6.5
	D19	27	0.184	18.1	0.135	13.7
	D3	11	0.274	25.2	0.212	21.9
Wilcoxon test p-value 0.333						
			3 bands		10 bands	
Species	Stratum	Observations	P	CV %	P	CV %
fulmar	A12	19	0.386	19.8	0.437	20.9
	A3	198	0.308	4	0.321	11.5
	A4	98	0.249	8.6	0.261	7.3
	A8	27	0.306	13.9	0.255	31.5
	B7	13	0.226	24.3	0.217	13
	D5	12	0.259	24.4	0.249	20.8
	D6	12	0.401	25.2	0.357	20.9
Wilcoxon test p-value 0.866						
			3 bands		10 bands	
Species	Stratum	Observations	P	CV %	P	CV %
black-backed gull	A5	93	0.27	8.7	0.266	7.5
	B10	24	0.239	17.6	0.28	20.8
	B11	51	0.288	7	0.282	21.1
	B12	35	0.249	14.4	0.337	20.8
	B2	71	0.239	10.2	0.25	8.6
	B3	19	0.303	15.9	0.34	20.7
	B8	10	0.165	30.4	0.122	24.2
	B9	31	0.267	24.4	0.316	30.8
	C1	26	0.277	16.4	0.372	27.3
	C10	19	0.276	19.1	0.245	16.6
	C11	44	0.326	27.2	0.372	20.9
	C12	16	0.228	21.8	0.217	18.2
	C2	131	0.327	15.8	0.328	14.1
	C3	69	0.264	10.1	0.288	8.7
	C9	44	0.281	12.6	0.285	22.4
	D12	47	0.251	12.4	0.303	6.8
	D15	97	0.241	8.7	0.252	21.5
	D16	24	0.283	17	0.299	14.7
Wilcoxon test p-value 0.306						

			3 bands		10 bands	
Species	Stratum	Observations	P	CV %	P	CV %
kittiwake	A2	17	0.373	20.7	0.446	33.8
	A3	27	0.27	26.3	0.272	29.7
	B10	19	0.243	19.7	0.218	11
	B8	35	0.288	8.4	0.21	36.1
	B9	51	0.268	11.7	0.271	10
	C8	20	0.298	18.5	0.302	16
	D13	18	0.239	20.3	0.197	17.2
	D17	20	0.208	20.2	0.255	16.1
	D19	18	0.327	19.6	0.28	17

Wilcoxon test p-value 0.953

			3 bands		10 bands	
Species	Stratum	Observations	P	CV %	P	CV %
auks	A10	66	0.262	16	0.264	5.4
	A11	80	0.308	6.3	0.296	11.9
	A12	25	0.27	27.3	0.276	46.1
	A3	227	0.303	9.9	0.282	9.8
	A4	228	0.331	5.1	0.315	6.3
	A5	22	0.181	20.2	0.195	15.6
	A6	62	0.272	10.6	0.263	9.15
	A/	90	0.188	9.9	0.2	18.1
	A8	43	0.204	13.9	0.191	11.2
	B10	60	0.264	1.9	0.281	19.5
	B2	11	0.231	26.2	0.237	21.8
	B3	21	0.187	20.5	0.171	16.1
	B4	95	0.249	8.7	0.229	7.4
	B5	81	0.243	52.1	0.221	5.5
	B6	207	0.295	4.1	0.272	11.2
	B7	164	0.367	6.9	0.349	12.5
	B8	163	0.338	5.4	0.369	16.1
	B9	197	0.299	4.6	0.309	9.8
	C12	10	0.247	27	0.242	22.9
	C3	10	0.194	29.3	0.222	22.9
	C4	49	0.229	12.5	0.212	10.4
	C5	272	0.421	12	0.409	9.5
	C6	119	0.341	6.5	0.362	12.3
	C7	45	0.191	13.9	0.174	11
	C8	131	0.328	6.5	0.284	13
	D1	165	0.305	5.5	0.313	19.34
	D10	228	0.282	9.2	0.27	9.35
	D11	121	0.25	7.7	0.292	5.1
	D12	65	0.263	15.9	0.255	20.3
	D13	37	0.288	8.2	0.246	5.1
	D16	88	0.303	8.8	0.354	17
	D17	223	0.286	3.2	0.295	17.5
	D18	68	0.269	4.7	0.285	32.8
	D19	290	0.237	5.1	0.224	3.8
	D2	119	0.236	16.6	0.228	10
	D3	221	0.294	11.8	0.293	11
	D4	131	0.316	5.7	0.342	20.9
	D6	189	0.3	6	0.354	11.6
	D7	190	0.239	12.8	0.223	12
	D8	120	0.218	23.7	0.214	15.7
	D9	26	0.245	16.8	0.244	11

Wilcoxon test p-value 0282

4 Discussion

4.1. Power analyses

4.1.1. General

The power analyses carried out as part of this study make it possible to determine the probability of detecting change in bird numbers, thus the number of birds displaced from a given wind farm. Although it is possible that some species, such as cormorant (*Phalacrocorax carbo*) could also be attracted to the wind farm by the presence of suitable structures for resting, or that colonisation of foundations by sessile organisms and flora, could eventually lead to improved bird feeding conditions (Leonhard and Pedersen, 2004), such behaviour is of less concern and consequently lower emphasis is placed on the possibility of attraction in this study. Furthermore, since the statistical power of detecting a given increase in numbers is unlikely to differ substantially from that of detecting a corresponding decrease, the results presented could equally be applied to gain good insight into the likelihood of detecting increases.

In providing an evidence-based assessment of whether existing aerial survey programmes are suitable for determining whether wind farms might result in changes in bird numbers, it has been necessary to make some assumptions. Perhaps the most critical of these was the assumption that the variance to mean ratio remains constant across all sites. This assumption was necessary because continuous and suitably long time-series of data were only available for a small number of sites and thus calculating variance to mean ratios separately for each site would limit the geographical scope of the study considerably and it would be unlikely that the results would be indicative of the full spectrum of sites. For the limited number of sites for which this assumption could be tested, it appears that it is a fairly valid assumption. Comparing actual variance to mean ratios for sites where more than four years of data are available to those predicted by assuming a constant variance to mean ratio across all sites, suggests that the variability in count data may have been slightly over-estimated and that, with the exception of sandwich tern, the variance to mean ratio is generally higher when mean counts are greater. However it is uncertain as to whether these are real trends or merely artefacts arising due to the small number of potentially unrepresentative sites used to make this assessment.

To determine the value of existing aerial baseline data for identifying changes in the numbers of birds, power analyses were carried out using data from a series of representative species: red-throated diver, common scoter, sandwich terns and black-backed gulls. Carrying out the analyses for a variety of species is essential as population dynamics (their numbers, spatial and temporal distribution) vary with species and as such both the mean and variance in numbers will thus also be species-specific. We opted for these species as all occur in nationally and / or internationally important numbers in the shallow offshore waters most favoured by wind farm developers and all, except black-backed gulls, are likely to be adversely displaced by wind farm construction and operation (Kaiser 2002; Petersen *et al.* 2005). For all of the species there are some issues associated with identification. The ways in which we attempted to circumvent identification problems are described in the methods, but throughout the study it is necessary to bare in mind the limitations imposed by the difficulty of identifying birds from a rapidly moving aircraft.

In order to limit the potential sources of variability in the baseline data we decided to use the raw data without correction for distance bias. Thus, the power analysis had to be made assuming detection probability does not vary greatly with the number of birds present. This assumption may not hold true entirely as large clusters of birds are likely to be more easily detected and this potential source of bias should be borne in mind when interpreting these results. Detection may also differ under different weather conditions and with the personal skill of the observer (neither of which were recorded in the data we had available for this study). Whilst there is no reason to suppose that these factors may introduce bias, they may account for some of the variability within the data and in future should be recorded. However, in general the DTI aerial surveys were carried out by a small number of trained observers, which would have reduced the impact of observer variability on the numbers of birds counted. The degree to which varying weather conditions have impacted on the numbers of birds observed is unknown, but may have been reduced due to the selection of relatively calm weather conditions during most aerial surveys.

Since in most instances the statistical power of detecting changes in bird numbers was fairly low, the results are thus presented with higher values of significance than would be typical of scientific studies in general (e.g. a significance of 0.2 rather than 0.05 is used). Relaxation of the level of significance results in a higher probability that an incorrect conclusion of "no effect" is accepted, but may improve the ability of a reasonable survey effort to detect change (Innogy 2003). Nevertheless full results with higher levels of statistical significance are presented in Annex 1. The effects of specific factors on the statistical power of detecting specified declines in numbers are now discussed.

4.1.2. Effects of spatial scale

For any given decline, statistical power and thus the likelihood of detecting changes increases with scale of analysis. This is primarily because larger areas support greater numbers of birds and it is easier to detect changes when more birds are present. However, wind farm footprint areas are usually fairly small in size, sometimes as small as 4 km^2, although often typically around 10 km x 10 km when the buffer area is included (Banks *et al.* 2005). Consequently, the decline due to a wind farm is likely to be relatively localised and unlikely to manifest itself throughout larger survey blocks. As such, a 50% decline in a 50 km x 50 km site is not directly comparable to a 50% decline in a 10 km x 10 km site as the former is unlikely to occur solely due to the presence of a wind farm. Adjusting the declines in larger survey blocks so that they are equivalent to declines in smaller survey area reveals that the statistical power generally increases as the area selected for analysis decreases.

We recommend therefore that the spatial scale used for analysis should be equivalent to that of the anticipated wind farm footprint and buffer area. In selecting a larger spatial scale for analysis there is a risk that changes in numbers would not occur throughout the surveyed area and that the likelihood of detecting changes due to the wind farms would decrease. In selecting a smaller area for analysis, it is likely that bird numbers within the area would be lower than if a larger area was used, and again, the probability of detecting changes in numbers would decrease. Aerial surveys allow flexibility in the choice of scale for analysis as data are recorded at high resolution. However, the resolution is constrained by the fact that detection declines with distance. As such, using the current survey methodology, spatial resolution mapping down to 2 km x 2 km is the minimum that can be achieved without biasing results and is also the minimum achievable without risking double counting (Camphuysen *et al.* 2004).

4.1.3. Effects of survey intensity

Since aerial surveys are relatively expensive, we compared the statistical power of conducting one aerial survey in every month throughout the year with that of conducting surveys only during the six months in which the species in question were most common: January to March and October to December for red-throated diver and common scoter and April to September for sandwich tern. Black-backed gulls occur throughout the year and consequently no comparative analyses were undertaken for these species.

Using all months as opposed to just selected months results in a fairly large increase in statistical power: c. 8% for red-throated diver, c. 4% for common scoter and c. 12% for sandwich tern. As such, the likelihood of detecting a 10% decline with surveys conducted in every month is broadly similar to the likelihood of detecting a 25% decline in red-throated diver and common scoter numbers and is greater than the likelihood of detecting a 50% decline in sandwich tern if surveys are conducted in only six months of the year. This would suggest that a significant number occur outside the six-month period in which they are most abundant. Consequently we condone extending the duration of the survey to encompass all months in which target species occur in significant numbers and if necessary, throughout the year. The latter may be necessary to ensure that changes in numbers of a species occurring both in summer and winter can be detected.

4.1.4. Effects of survey duration

Contrary to expectation, simulating data to increase the length of time over which surveys were conducted did not substantially increase the power of detecting changes in numbers. Adopting a survey protocol that entails surveys in every month for 10 years prior to construction followed by surveys in every month for ten years after construction as opposed to monthly surveys for just two years before and two years after construction only increased the likelihood of detecting changes by 5 - 15%. A five years pre- and post-construction survey only increased power by 5 - 10% over two year pre- and post surveys. In general however the chances of detecting a 10% change using a total of twenty years of survey is broadly equivalent to the likelihood of detecting a 50% change using a total of four years survey. It is also noteworthy that for the selected time-periods, an additional one-year of pre- and one year of post-construction survey results in disproportionate increase in the likelihood of detecting changes, although the exact time period is itself species dependent. In the case of red-throated diver, the greatest proportional improvement is achieved by extending the survey from four years to six years, for the remainder of the species the most significant improvements would be achieved if surveys were extended from six to eight years.

Given the additional cost of conducting surveys for longer periods, doing so may not be the most cost-effective means of attempting to detect changes in numbers. Nevertheless, there may be biological as opposed to statistical reasons for undertaking surveys for a greater length of time. It is possible that, after an initial period of time birds become habituated to wind farm related disturbance and consequently return after a period of initial displacement. Without a sufficiently long survey period any such behaviour would not be detected.

4.1.5. Effects of including a reference area

The need to distinguish changes in bird abundance due to the wind farm development itself, rather than any potential confounding factors such as long-term changes in population size or distribution shifts in response to changing climate and weather, necessitates the use of a reference area. Any changes occurring within the wind farm footprint and buffer areas can then be compared to changes occurring within the reference area. This BACI design has its limitations, in part because natural processes can also induce location-specific changes (Underwood 1994), changes that may be particularly difficult to predict in marine environments, and in part because birds displaced from the wind farm may travel to neighbouring areas including the reference area itself. Whilst this will increase the likelihood of detecting changes, it may give a false impression of the magnitude of the change in relation to any changes independent of the wind farm. Nevertheless comparisons of changes in numbers to those occurring within such a reference area is a considerable improvement of not doing so. For this reason, we opted to investigate the impact of including a reference area on the statistical power of detecting changes. In general the inclusion of a reference area decreases the probability that large declines in birds will be detected but increases the likelihood of detecting small changes. Using a larger reference area as opposed to a small area, makes little difference, with the minimal observed differences being largely species specific. Consequently we recommend that comparisons to a reference area should be made, particularly if relatively small changes are expected. The size and location of this area should not be governed by the extent to which it would increase capacity to detect changes in bird numbers, but should give due consideration to the biology and behaviour of the species in question. Indeed, an assessment of changes within every 10 km x 10 km grid cell with a larger aerial survey block (i.e. the area that can be reasonably covered in one day) is likely to give much greater insight into spatial patterns in displacement than just comparing changes in numbers in the footprint and buffer area to a single reference area. The additional costs of doing so a likely to be minimal, only requiring a small increase in the time taken to analyse the data.

4.1.6. Effects of mean and peak counts

By far the most important determinant of statistical power appears to be the mean number of birds recorded on a particular site. This finding is reassuring, as it would suggest that the more important a site is for a particular species, the more likely it is that change would be detected. Nevertheless, relative to the number of birds that might be considered important, the likelihood of detecting changes varies greatly between species and is, for more aggregated species like notably common scoter and red-throated diver, rather small, while for less aggregated species like sandwich tern and black-backed gulls it is adequate. Even if one assumes a large (50%) decline and relaxes power of detecting change to 80% and statistical significance to 0.2, the peak number that would have to be present prior to wind-farm construction, before changes could be detected would be in excess of 1% of the national populations of red-throated diver and common scoter, the thresholds used to determine whether a site is nationally important. Additionally, the figures presented in Table 3.2.1 refer to raw counts. In reality only around one-fifth to a-third of all birds are actually detected, so actual numbers would have to be in three to five times higher the peak counts listed in the table before changes could be detected. Of even more concern, the threshold for acceptable declines in numbers of birds due to wind farm-induced habitat displacement would be considerably lower than 50%, and if statistical assumptions are tightened so that they are in line with most scientific studies: i.e. statistical power must exceed 95% power and statistical significance was be less than 0.05, the numbers of birds that would be needed to detect change surpass international thresholds for all species.

4.2. Covariate analyses

In general, one would expect that incorporating meteorological, topographic and hydro-dynamic covariates into models would increase the likelihood of detecting changes as these are known to reduce the unexplained variability in the count data. However, by using the approach adopted in this study, whereby each site is considered on an individual basis, incorporating spatial covariates into the model that do not vary temporally cannot explain part of the variance in counts as such variance would be explained by the incorporation of a site effect into the model in any case. In order to improve the statistical power through the use of covariates, it would be necessary to incorporate hydro-dynamic covariates. As the incorporation of dynamic (i.e. temporally variable) covariates was not possible within the time-frame of this study, we chose to investigate the effects these might have by comparing whether there is a reduction in variance to mean ratios when static covariates are used.

The incorporation of static covariates does indeed help to explain much of the variance in count data and there is thus every reason to expect that incorporating hydrodynamic data would explain a greater proportion of the variance and thus enhance the likelihood of detecting changes in numbers. Indeed, given the limited improvements in statistical power obtained by other methods, doing so would appear to be the most advisable way of increasing the statistical power to an acceptable level for the more aggregated species.

4.3. Analysis of the importance of resolution in collected distances

The variability of estimated densities or numbers of birds recorded on line transect surveys can be partitioned into three main components: variance due to detection probability, encounter rate and cluster size. Of these, encounter rate generally determines most of the total variance of the estimated density, and therefore improvements of the quality of the estimated density are typically made by increasing the number of samples (transect lines). However, since the design of the standard aerial line transect design used by the DTI scheme differs markedly from most standard line transect designs by using a minimum number of perpendicular distance bands it was tested whether the introduction of a higher resolution in collected distances would increase detection probability. The results of the comparison of estimated detection probabilities between distance analysis designs with three and 10 bands are quite clear, and demonstrate that given the low detection probability at short distances, the flat shoulder of observations within 100 m distance and the sharp decline in detections between 100m and 200m the precision of estimated densities from the aerial surveys will only be increased marginally by the application of shorter distance categories or collection of real rather than grouped distances.

It is recommended to carry out an ongoing assessment of the data collected during the DTI surveys, especially with a view to improve the possibility to better apply detection functions and increase the understanding of potential sampling errors which may bias the estimation of density. Although unbiased estimates of the density of birds are not a prerequisite for monitoring seabirds in relation to windfarms, population estimates produced by the surveys play an important role in the determination of the significance of the estimated impacts.

4.4. Optimal surveying strategy and improvements to future surveys

Assuming little limitations to cost, current survey methods would be improved by maximising the number of counts within the desired time-period and by extending this time period for as long as possible. Satisfactory statistical power could be achieved as follows:

- Months - Survey at least once every month throughout the year as opposed to for just six months as this leads to a significant improvement in power (increasing it by a factor of 1.5 in the case of sandwich tern).
- Time period - Survey for four years prior to construction and for four years after construction. In the case of red-throated diver, the greatest proportional improvement is achieved by extending the survey from two years to three years pre- and post-construction and for the remainder of the species the most significant improvements is achieved if surveys were extended from three to four years pre- and post- construction.
- Scale - always analyse data at a scale equivalent to that of the wind farm footprint and buffer area as this maximises the chances of detecting changes in numbers.
- Reference area - since the size of the reference area has little predictable influence on statistical power, the reference area should be designed taking due consideration of the biology and behaviour of the species in question rather than being governed by the need to detect changes.
- Covariates - if possible, collect hydro-dynamic covariates as this is likely to improve the power of detecting changes.

Assuming that the cost of aerial surveys is a factor in determining design, optimal survey design could be achieved as follows:

- Months - Survey at least once every month throughout the year as some species are likely to be most numerous in winter and others in summer. If this is not the case, the improvement in statistical power is not sufficient to justify surveying throughout the year. Since the probability of detecting change is most influenced by mean numbers, greater improvements in statistical power would be achieved by concentrating survey effort at the time of year in which the species in question are most abundant.
- Time period - Survey for as long as one expects changes to manifest themselves over. Improvements in power obtained by surveying for longer periods do not justify the additional costs. Improvements in power would be obtained more cost-effectively by surveying more intensively over a shorter period.
- Scale - always analyse data at a scale equivalent to that of the wind farm footprint and buffer area as this maximises the chances of detecting changes in numbers.
- Reference area - since the size of the reference area has little predictable influence on statistical power the design reference area by taking due consideration of the biology and behaviour of the species in question rather than being governed by the need to detect changes.
- Covariates - if possible, collect hydro-dynamic covariates as this is likely to improve the power of detecting changes.

Taking the pragmatic approach outlined above, the best power that can be obtained for specified declines, assuming a statistical significances of 0.2 is given in Table 4.2.

Table 4.2. Highest statistical power than can be attained by adopting a pragmatic approach for specified declines.

Species	**50% decline**	**25% decline**	**10% decline**
red-throated diver	51.12%	51.42%	49.18%
common scoter	47.01%	43.67%	40.25%
black-backed gulls	61.30%	55.35%	56.65%
sandwich tern	66.25%	59.44%	60.14%

The collection of synoptic dynamic covariate data and the inclusion of such data into analysis is likely to be one of the best means of improving the statistical power of detecting changes in bird numbers. At present, however, the exact effect this would have on the likelihood of detecting changes remains untested, except through the proxy methods presented in this study. There are many possible ways to obtain synoptic covariate data, but the most feasible in the context of this study would be to create a hydro-dynamic model with high spatial and temporal resolution and parameterise and calibrate this model using local hydrographic measurements. Surface temperature and salinity can be sampled from aeroplanes using radar and profiles of temperature, salinity and currents obtained using disposable profilers. This option is expensive and only provides data on surface patterns, and requires intensive post-processing chains to produce meaningful explanatory variables of bird abundance. Furthermore, disposable profilers, dumped into the ocean, could cause damage to the environment. An alternative approach is to obtain oceanographic data using concurrent ship-based surveys. One major problem with a concurrent ship-based program is that boats cannot travel at the same speed as aircraft and consequently it is impossible to obtain oceanographic data that are truly concurrent with bird counts. Given the rapid variability in oceanographic parameters (e.g. change in tidal-induced currents and location of related fronts), survey-based collections of oceanographic data are deemed insufficient to resolve the changes taking place in the water column during a survey (Dippner 1993) Satellite remote sensing, especially NOAA AVHRR SST, can be obtained several times per day, but due to cloud cover synoptic sampling is severely constrained.

However, with the development of operational hydrodynamic forecast services and the increasing use of hydrodynamic models in local environmental studies it is now possible to make reliable fine-scale (e.g. 500 m horizontal, 2 m vertical, half-hour temporal resolution) predictions of the three-dimensional current and density structures of the water column and water levels in any of the development regions.

4.5. Conclusions

Current aerial survey methods provide adequate means of detecting changes in the numbers of dispersed species like sandwich terns and black-backed gulls and are consequently likely to detect changes for most species that are not prone to large natural fluctuations in numbers between years. For those species, such as red-throated diver and common scoter which are aggregated and whose numbers are prone to a greater degree of inter-annual fluctuation, existing aerial survey methods only provide restrained means of detecting changes in regions in which these species are particularly abundant. Whilst increasing the frequency or duration of aerial surveys provides one means of increasing the likelihood of detecting changes in bird numbers, incorporating hydro-dynamic covariate data into analysis is likely to provide a more cost-effective means of achieving this.

Acknowledgements

This project was COWRIE funded under DISP-03-2006. We thank Andrew Prior (JNCC) for managing this project, Andy Webb (JNCC), Brehon Frost (RPS) and especially Philip Bloor (DTI) for providing us with aerial survey data. Alex Banks (BTO), Bill Sanderson (CCW) and Sian Whitehead (CCW) kindly provided additional scoter data. Thanks also to Steve Freeman (BTO), Niall Burton (BTO and Nick Carter (BTO) and Martin Laczny (Biola) for commenting on aspects of this report.

5 Annex 1 - Details of power analysis

5.1. Baseline results

Table 5.1.1 Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction, all months.

% decline	**P**	**red-throated diver**		**common scoter**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.70%	37.31%	47.01%	9.86%	66.25%	15.29%
50	0.1	45.64%	34.96%	46.63%	9.79%	61.94%	14.29%
50	0.05	42.03%	32.20%	41.79%	8.77%	57.50%	13.27%
25	0.2	45.28%	34.69%	43.67%	9.16%	59.17%	13.65%
25	0.1	40.25%	30.83%	36.96%	7.76%	52.08%	12.02%
25	0.05	36.24%	27.76%	34.42%	7.22%	49.01%	11.31%
10	0.2	40.21%	30.80%	40.25%	8.45%	56.39%	13.01%
10	0.1	37.64%	28.83%	33.81%	7.09%	48.61%	11.22%
10	0.05	32.18%	24.65%	28.85%	6.05%	41.00%	9.46%

5.2. Effects of spatial scale

Table 5.2.1. Statistical power for 50%, 25, and 10% declines using a spatial scale of 5 km x 5 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction.

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	45.25%	30.08%	44.48%	5.99%	56.86%	38.92%	55.65%	8.45%
50	0.1	41.24%	27.41%	40.55%	5.46%	49.12%	33.62%	46.67%	7.09%
50	0.05	38.00%	25.26%	34.91%	4.70%	41.99%	28.74%	43.96%	6.68%
25	0.2	41.53%	27.60%	37.93%	5.11%	46.87%	32.08%	46.10%	7.00%
25	0.1	37.03%	24.61%	35.92%	4.84%	37.84%	25.90%	40.92%	6.21%
25	0.05	33.15%	22.03%	30.87%	4.16%	29.81%	20.40%	35.06%	5.32%
10	0.2	39.54%	26.28%	35.94%	4.84%	41.49%	28.40%	47.66%	7.24%
10	0.1	33.79%	22.46%	30.50%	4.11%	32.71%	22.39%	38.64%	5.87%
10	0.05	30.24%	20.10%	24.36%	3.28%	26.28%	17.99%	32.75%	4.97%

Table 5.2.2. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction.

% decline	P	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.70%	37.31%	47.01%	9.86%	61.30%	49.90%	66.25%	15.29%
50	0.1	45.64%	34.96%	46.63%	9.79%	46.10%	37.53%	61.94%	14.29%
50	0.05	42.03%	32.20%	41.79%	8.77%	45.67%	37.18%	57.50%	13.27%
25	0.2	45.28%	34.69%	43.67%	9.16%	49.10%	39.97%	59.17%	13.65%
25	0.1	40.25%	30.83%	36.96%	7.76%	40.87%	33.27%	52.08%	12.02%
25	0.05	36.24%	27.76%	34.42%	7.22%	32.63%	26.56%	49.01%	11.31%
10	0.2	40.21%	30.80%	40.25%	8.45%	42.20%	34.36%	56.39%	13.01%
10	0.1	37.64%	28.83%	33.81%	7.09%	35.18%	28.64%	48.61%	11.22%
10	0.05	32.18%	24.65%	28.85%	6.05%	29.59%	24.09%	41.00%	9.46%

Table 5.2.3. Statistical power for 50%, 25, and 10% declines using a spatial scale of 50 km x 50 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction.

% decline	P	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	52.50%	48.46%	56.84%	30.86%	70.00%	67.31%	66.67%	38.46%
50	0.1	50.00%	46.15%	56.00%	30.40%	68.33%	65.70%	56.67%	32.69%
50	0.05	45.00%	41.54%	52.22%	28.35%	54.17%	52.09%	62.00%	35.77%
25	0.2	51.82%	47.83%	56.67%	30.76%	59.57%	57.28%	61.33%	35.38%
25	0.1	41.25%	38.08%	50.00%	27.14%	48.80%	46.92%	48.67%	28.08%
25	0.05	40.00%	36.92%	48.42%	26.29%	42.92%	41.27%	45.33%	26.15%
10	0.2	48.26%	44.55%	46.19%	25.07%	54.58%	52.48%	60.00%	34.62%
10	0.1	39.20%	36.18%	45.50%	24.70%	44.58%	42.87%	52.67%	30.39%
10	0.05	37.20%	34.34%	48.75%	26.46%	40.87%	39.30%	43.57%	25.14%

Table 5.2.4. Statistical power for declines equivalent to 100% in a 2 km x 2 km block, for a range of spatial scales (p = 0.2).

Scale (km)	Percentage decline	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
2.2 x 2.2	82.64%	49.41%	21.02%	52.28%	4.03%	65.00%	23.15%	45.85%	2.52%
2.5 x 2.5	64.00%	47.56%	17.17%	44.42%	3.23%	57.65%	22.78%	51.74%	3.16%
3 x 3	44.44%	45.70%	23.21%	41.91%	4.10%	53.11%	24.75%	53.48%	4.55%
4 x 4	25.00%	41.70%	25.41%	39.41%	4.40%	44.82%	27.20%	44.39%	5.40%
5 x 5	16.00%	40.85%	27.15%	37.35%	5.03%	42.98%	29.42%	47.27%	7.18%
6 x 6	11.11%	39.61%	27.92%	37.69%	4.91%	42.82%	30.97%	46.84%	8.21%
7 x 7	8.16%	40.24%	28.92%	37.48%	6.40%	42.83%	32.14%	48.87%	9.88%
8 x 8	6.25%	39.88%	29.87%	37.66%	7.13%	40.08%	32.05%	48.74%	10.20%
9 x 9	4.94%	40.58%	29.97%	37.72%	7.23%	42.29%	34.42%	52.00%	12.61%
10 x 10	4.00%	41.34%	31.67%	38.02%	7.98%	42.48%	34.58%	54.44%	12.56%

5.3. Effects of survey intensity

Table 5.3.1 Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction, all months.

% decline	P	red-throated diver		common scoter		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.70%	37.31%	47.01%	9.86%	66.25%	15.29%
50	0.1	45.64%	34.96%	46.63%	9.79%	61.94%	14.29%
50	0.05	42.03%	32.20%	41.79%	8.77%	57.50%	13.27%
25	0.2	45.28%	34.69%	43.67%	9.16%	59.17%	13.65%
25	0.1	40.25%	30.83%	36.96%	7.76%	52.08%	12.02%
25	0.05	36.24%	27.76%	34.42%	7.22%	49.01%	11.31%
10	0.2	40.21%	30.80%	40.25%	8.45%	56.39%	13.01%
10	0.1	37.64%	28.83%	33.81%	7.09%	48.61%	11.22%
10	0.05	32.18%	24.65%	28.85%	6.05%	41.00%	9.46%

Table 5.3.2 Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction, six months.

% decline	P	red-throated diver		common scoter		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	46.87%	35.90%	43.78%	9.19%	48.86%	11.28%
50	0.1	41.61%	31.87%	34.31%	7.20%	40.43%	9.33%
50	0.05	36.47%	27.94%	30.00%	6.30%	37.71%	8.70%
25	0.2	39.69%	30.40%	34.59%	7.26%	42.71%	9.86%
25	0.1	35.02%	26.83%	30.82%	6.47%	34.35%	7.93%
25	0.05	31.89%	24.43%	27.81%	5.84%	28.96%	6.68%
10	0.2	37.56%	28.77%	32.97%	6.92%	40.28%	9.30%
10	0.1	33.53%	25.68%	27.81%	5.84%	34.14%	7.88%
10	0.05	31.32%	23.99%	25.65%	5.38%	30.15%	6.96%

5.4. Effects of survey duration

Table 5.4.1. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction.

% decline	P	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.70%	37.31%	47.01%	9.86%	61.30%	49.90%	66.25%	15.29%
50	0.1	45.64%	34.96%	46.63%	9.79%	46.10%	37.53%	61.94%	14.29%
50	0.05	42.03%	32.20%	41.79%	8.77%	45.67%	37.18%	57.50%	13.27%
25	0.2	45.28%	34.69%	43.67%	9.16%	49.10%	39.97%	59.17%	13.65%
25	0.1	40.25%	30.83%	36.96%	7.76%	40.87%	33.27%	52.08%	12.02%
25	0.05	36.24%	27.76%	34.42%	7.22%	32.63%	26.56%	49.01%	11.31%
10	0.2	40.21%	30.80%	40.25%	8.45%	42.20%	34.36%	56.39%	13.01%
10	0.1	37.64%	28.83%	33.81%	7.09%	35.18%	28.64%	48.61%	11.22%
10	0.05	32.18%	24.65%	28.85%	6.05%	29.59%	24.09%	41.00%	9.46%

Table 5.4.2. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 3-years pre-construction and 3-years post-construction.

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	53.29%	40.82%	48.94%	10.27%	61.80%	50.31%	67.36%	15.54%
50	0.1	48.32%	37.01%	48.25%	10.13%	56.30%	45.83%	59.58%	13.75%
50	0.05	46.67%	35.75%	41.75%	8.76%	48.59%	39.56%	55.14%	12.72%
25	0.2	49.54%	37.95%	43.58%	9.15%	50.04%	40.74%	59.58%	13.75%
25	0.1	44.17%	33.84%	38.02%	7.98%	42.09%	34.27%	49.58%	11.44%
25	0.05	41.75%	31.98%	36.34%	7.63%	32.40%	26.38%	44.58%	10.29%
10	0.2	46.10%	35.31%	41.98%	8.81%	43.18%	35.15%	52.50%	12.12%
10	0.1	40.88%	31.32%	35.37%	7.42%	35.35%	28.78%	45.00%	10.38%
10	0.05	35.35%	27.08%	29.18%	6.12%	27.60%	22.47%	42.78%	9.87%

Table 5.4.3. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 4-years pre-construction and 4-years post-construction.

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	57.03%	43.69%	53.53%	11.23%	67.25%	54.75%	71.94%	16.60%
50	0.1	51.37%	39.35%	50.74%	10.65%	59.03%	48.06%	67.08%	15.48%
50	0.05	47.91%	36.70%	44.57%	9.35%	51.10%	41.60%	63.75%	14.71%
25	0.2	50.50%	38.68%	45.91%	9.63%	53.15%	43.27%	64.31%	14.84%
25	0.1	45.18%	34.61%	39.55%	8.30%	43.36%	35.30%	56.94%	13.14%
25	0.05	42.21%	32.33%	37.53%	7.88%	33.24%	27.06%	50.42%	11.64%
10	0.2	48.31%	37.01%	41.26%	8.66%	44.10%	35.90%	59.72%	13.78%
10	0.1	40.08%	30.70%	35.73%	7.50%	34.88%	28.40%	52.36%	12.08%
10	0.05	35.72%	27.36%	36.54%	7.67%	27.36%	22.27%	45.97%	10.61%

Table 5.4.4. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 5-years pre-construction and 5-years post-construction.

% decline	P	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	57.46%	44.02%	53.45%	11.22%	69.72%	56.76%	71.94%	16.60%
50	0.1	53.94%	41.32%	50.00%	10.49%	61.36%	49.95%	69.39%	16.01%
50	0.05	50.72%	38.85%	45.70%	9.59%	54.07%	44.02%	65.69%	15.16%
25	0.2	51.20%	39.22%	46.86%	9.83%	54.96%	44.74%	62.22%	14.36%
25	0.1	48.66%	37.27%	44.42%	9.32%	46.72%	38.03%	57.92%	13.37%
25	0.05	43.12%	33.03%	40.91%	8.58%	38.20%	31.10%	49.31%	11.38%
10	0.2	49.44%	37.87%	43.22%	9.07%	47.51%	38.68%	59.31%	13.69%
10	0.1	41.88%	32.08%	38.35%	8.05%	37.02%	30.14%	54.03%	12.47%
10	0.05	38.70%	29.65%	34.66%	7.27%	29.96%	24.39%	45.28%	10.45%

Table 5.4.5. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 10-years pre-construction and 10-years post-construction.

% decline	P	red-throated diver		common scoter		black-backed gulls		sandwich tern	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	65.47%	50.15%	54.85%	11.51%	75.08%	61.12%	82.92%	19.14%
50	0.1	60.35%	46.23%	52.02%	10.92%	68.64%	55.88%	78.06%	18.01%
50	0.05	56.05%	42.94%	50.80%	10.66%	60.78%	49.48%	70.97%	16.38%
25	0.2	57.88%	44.34%	54.83%	11.51%	59.38%	48.34%	69.44%	16.02%
25	0.1	50.89%	38.98%	46.70%	9.80%	49.22%	40.07%	59.86%	13.81%
25	0.05	43.95%	33.67%	42.86%	8.99%	40.28%	32.79%	58.89%	13.59%
10	0.2	50.74%	38.87%	47.37%	9.94%	46.67%	37.99%	64.72%	14.94%
10	0.1	44.62%	34.18%	40.51%	8.50%	37.19%	30.28%	57.08%	13.17%
10	0.05	38.30%	29.34%	38.06%	7.99%	27.97%	22.77%	50.00%	11.54%

5.5. Effects of including a reference area

Table 5.5.1. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, for a range of significance levels (p = 0.05, 0.1 and 0.2). Time period = 2-years pre-construction and 2-years post-construction.

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.70%	37.31%	47.01%	9.86%	61.30%	49.90%	66.25%	15.29%
50	0.1	45.64%	34.96%	46.63%	9.79%	46.10%	37.53%	61.94%	14.29%
50	0.05	42.03%	32.20%	41.79%	8.77%	45.67%	37.18%	57.50%	13.27%
25	0.2	45.28%	34.69%	43.67%	9.16%	49.10%	39.97%	59.17%	13.65%
25	0.1	40.25%	30.83%	36.96%	7.76%	40.87%	33.27%	52.08%	12.02%
25	0.05	36.24%	27.76%	34.42%	7.22%	32.63%	26.56%	49.01%	11.31%
10	0.2	40.21%	30.80%	40.25%	8.45%	42.20%	34.36%	56.39%	13.01%
10	0.1	37.64%	28.83%	33.81%	7.09%	35.18%	28.64%	48.61%	11.22%
10	0.05	32.18%	24.65%	28.85%	6.05%	29.59%	24.09%	41.00%	9.46%

Table 5.5.2. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, with four adjacent 10 by 10 km squares to the north, south, east and west, for a range of significance levels (p = 0.05, 0.1 and 0.2).

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	48.08%	36.83%	35.76%	7.50%	55.55%	45.22%	58.06%	13.40%
50	0.1	44.47%	34.07%	32.56%	6.83%	48.75%	39.69%	49.03%	11.31%
50	0.05	40.34%	30.90%	25.68%	5.39%	40.56%	33.02%	45.00%	10.38%
25	0.2	46.63%	35.72%	37.67%	7.90%	53.79%	43.79%	58.47%	13.49%
25	0.1	42.53%	32.58%	30.12%	6.32%	45.47%	37.02%	48.61%	11.22%
25	0.05	40.13%	30.74%	25.26%	5.30%	39.88%	32.47%	41.39%	9.55%
10	0.2	45.86%	35.13%	36.79%	7.72%	53.23%	43.33%	56.67%	13.08%
10	0.1	42.31%	32.41%	32.03%	6.72%	46.32%	37.71%	49.31%	11.38%
10	0.05	39.19%	30.02%	24.94%	5.23%	36.71%	29.89%	40.42%	9.33%

Table 5.5.3. Statistical power for 50%, 25, and 10% declines using a spatial scale of 10 km x 10 km, with eight adjacent 10 by 10 km squares to the north, north-east, south, south-east, east north-west, west and south-west, for a range of significance levels (p = 0.05, 0.1 and 0.2).

% decline	**P**	**red-throated diver**		**common scoter**		**black-backed gulls**		**sandwich tern**	
		Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros	Non-zeros only	Including zeros
50	0.2	51.12%	39.16%	37.67%	7.90%	57.17%	46.54%	59.86%	13.81%
50	0.1	48.06%	36.82%	33.81%	7.09%	49.18%	40.04%	54.93%	12.68%
50	0.05	42.30%	32.40%	28.99%	6.08%	43.00%	35.01%	50.28%	11.60%
25	0.2	51.42%	39.39%	38.47%	8.07%	55.35%	45.06%	59.44%	13.72%
25	0.1	44.71%	34.25%	33.05%	6.94%	48.71%	39.65%	56.25%	12.98%
25	0.05	42.70%	32.71%	29.48%	6.19%	39.13%	31.86%	46.81%	10.80%
10	0.2	49.18%	37.67%	36.70%	7.70%	56.65%	46.12%	60.14%	13.88%
10	0.1	45.13%	34.57%	30.83%	6.47%	47.19%	38.42%	55.69%	12.85%
10	0.05	40.80%	31.25%	29.09%	6.10%	39.04%	31.78%	45.83%	10.58%

References

Akaike, H. (1976) An information criterion (AIC) *Mathematical Sciences* **14**, 5–9.

Baker, H., Stroud, D.A., Aebischer, N.J., Cranswick, P.A., Gregory, R.D., McSorley, C.A., Noble, D.G. and Rehfisch, M.M. (2005) Population estimates of birds in Great Britain and the United Kingdom. *British Birds* **99**, 25-44.

Banks, A.N., Burton, N.H.K., Austin, G.E., Carter, N., Chamberlain, D.N., Holt, C., Rehfisch, M.M., Wakefield, E. and Gill, P. (2005) *The potential effects on birds of the Greater Gabbard Offshore Wind Farm: Report for February 2004 to March 2005*. BTO Research Report No. 419, Thetford.

Banks, A.N., Maclean, I.M.D., Burton, N.H.K., Austin, G.E., Carter, N., Chamberlain, D.N., Holt, C., Rehfisch, M.M., Pinder, S., Batty, A., Wakefield, E. and Gill, P. (2006) *The potential effects on birds of the Greater Gabbard Offshore Wind Farm: Report for February 2004 to April 2006*. BTO Research Report No. 440, Thetford.

Blomdahl, A., Breife, B. and Holmstrom, N. (2003) *Flight identification of European Seabirds*. Helm, London.

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

Camphuysen, C.J., Fox, A.D., Leopold, M.F. and Petersen, I.K. (2004) *Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the U.K*. A report commissioned by COWRIE. Royal Netherlands Institute for Sea Research.

Carter, I.C., Williams, J.M., Webb, A. and Tasker, M.L. (1993) *Seabird concentrations in the North Sea: an atlas of vulnerability to surface pollutants.* JNCC, Aberdeen.

Chamberlain, D.E., Rehfisch, M.R., Fox. T., Desholm, M. and Anthony, S.J. (2006) The effect of avoidance rates on bird mortality predictions made by wind turbine collision risk models. *Ibis*, **148**, 198-202.

Christensen, T.K and Hounisen, J.P. (2005) *Investigations of migratory birds during operation of Horns Rev offshore wind farm*. National Environmental Research Institute, Ministry of the Environment, Denmark.

Cohen, J. (1988) *Statistical Power Analysis for Behavioural Sciences.* 2nd Edition. Lawrance Erblaum Associates, Hillsdale, New Jersey.

Collier, M, Banks, A., Austin, G., Girling, T., Hearn, R. and Musgrove, A. (2005) *The Wetland Bird Survey 2003-04: Wildfowl and Wader Counts*. BTO/WWT/RSPB/JNCC.

Dean, B.J., Webb, A., McSorley, C.A. and Reid, J.B. (2003) *Aerial surveys of inshore areas for wintering seaducks, divers and grebes: 200/01 and 2001/02*. JNCC Report No 333, Peterborough.

Desholm, M. and Kahlert, J. (2005) Avian collision risk at an offshore wind farm. *Biology Letters* **13**, 1-4.

Desholm, M., Fox, T. and Beasley, P. (2005) Best practice guidance for the use of remote techniques for observing bird behaviour in relation to offshore windfarms. A report commissioned by COWRIE. COWRIE – REMOTE-05-2004.

Diederichs, A., Nehls, G. and Petersen, I.K. (2002): Flugzeugzählungen zur großflächigen Erfassung von Seevögeln und marinen Säugern als Grundlage für Umweltverträglichkeitsstudien im Offshorebereich. - *Seevögel* **23**, 38-46.

Dippner, J. (1993) A frontal-resolving model for the German Bight. *Continental Shelf Research* **13**, 49-66.

Drewitt, A.L. and Langston, R.H.W. (2006) Assessing the impacts of wind farms on birds. *Ibis*, **148**, 29-42.

Durinck, J., Skov, H. and Andell., P. (1993) Seabird distribution and numbers in selected offshore parts of the Baltic Sea, winter (1992) *Ornis Svecica*, 3: 11-26.

Durinck, J., Skov, H., Jensen, F.P. and Pihl, S. (1994) *Important marine areas for wintering birds in the Baltic Sea.* Ornis Conult report to the European Commission.

Exo, K.M., Hüppop, O. and Garthe, S. (2003) Birds and offshore wind farms: a hot topic in marine ecology. *Wader Study Group Bulletin* **100**, 50-53.

Federal Maritime and Hydrographic Agency (2003) *Standards for Environmental Impact Assessments of Offshore Wind Turbines in the Marine Environment (Status: 25. February 2003).* Federal Maritime and Hydrographic Agency, Hamburg, Germany. Available from: http://www.bsh.de/en/Marine%20uses/Industry/Wind%20farms/standard_environmental.pdf.

Fretwell, S.D. and Lucas, H.L. (1970) On territorial behaviour and other factors influencing habitat distributions in birds. *Acta Biotheoretica* **19**, 16-36.

Garthe, S. and Hüppop, O. (2004) Scaling possible adverse effects of marine wind farms on seabirds: developing and applying a vulnerability index. *Journal of Applied Ecology* **41**, 724-734.

Hötker, H., Thomsen, K.-M. and Köster, H. (2004) *Auswirkungen regenerativer Energiegewinnung auf die biologische Vielfalt am Beispiel der Vögel und der Fledermäuse – Fakten, Wissenslücken, Anforderungen an die Forschung, ornithologische Kriterien zum Ausbau von regenerativen Energiegewinnungsformen. Endbericht 2004.* Michael-Otto-Institut im NABU, Bergenhusen, Germany. Available from: http://bergenhusen.nabu.de/bericht/VoegelRegEnergien.pdf.

Innogy (2003) *North Hoyle Offshore wind farm baseline monitoring report*. NH/MonR/Rev 1.

Joensen, A. H. (1968) Wildlife counts in Denmark in November 1967 and January 1968. *Danish Review of Game Biology* **5**, 1-72.

Joensen, A. H. (1973) Moult migration and wing-feather moult of seaducks in Denmark. *Danish Review of Game Biology* **8**, 1-43.

Joensen, A. H. (1974) Waterfowl populations in Denmark 1965-73. *Danish Review of Game Biology* **9**, 1-206.

JNCC (2004) *Guidance on Offshore Windfarm Development - extract of draft guidance on bird survey techniques*. JNCC, Peterborough.

Kahlert, J., Desholm, M., Clausager, I. and Petersen, I.K. (2000) *Environmental impact assessment of an offshore wind park at Rødsand: Technical report on birds*. National Environmental Research Institute, Ministry of the Environment, Denmark.

Kahlert, J., Peterson, I.K., Desholm, M. and Therkildsen, O. (2004) *Investigations of birds during operation of Nysted offshore wind farm at Rødsand: Results and conclusions, 2004*. National Environmental Research Institute, Ministry of the Environment, Denmark.

Kaiser, M. (2002) *Predicting the displacement of common scoter* Melanitta nigra *from benthic feeding areas due to offshore windfarms*. A report commissioned by Cowrie. COWRIE-BEN-2002-03.

Laursen, (1989) Estimates of sea duck winter populations in the Western Palaearctic. *Danish Review of Game Biology* **13**, 1-22.

Laursen, K., Pihl, S., Durinck, J., Hansen, M., Skov, H., Frikke, J. and Danielsen, F. (1997) Numbers and distribution of waterbirds in Denmark 1987-89. *Danish Review of Game Biology* **15**, 1-181.

Leonhard, S.B. and Pedersen, J.P. (2004) *Hard Bottom Substrate Monitoring*. Horns Rev Offshore Wind Farm. Annual Status Report 2003. Bioconsult.

Mackey, M. and Giménez, D.P. (2000) *SEA678 Data Report for Offshore Seabird Populations*. DTI report.

Merck, T. and von Nordheim, H. (2002) Technische Eingriffe in marine Lebensräume - Workshop des Bundesamtes für Naturschutz, International Naturschutzakademic Insel Vilm, 27-29 Oktober 1999. Bonn-Bad Godesberg, Germany.

Percival, S.M., Band, B. and Leeming, T. (1999) *Assessing the ornithological effects of wind farms: developing a standard methodology*. Proceedings of the 21st British Wind Energy Association Conference 1999.

Petersen, I.K. (2005) *Bird numbers and distribution in the Horns Rev offshore wind farm area. Annual status report 2004*. Report commissioned by Elsam Engineering A/S 2005. National Environmental. Research Institute, Rønde, Denmark.

Petersen, I.K., Clausager, I. and Christensen, T.J. (2004) *Bird numbers and distribution in the Horns Rev offshore wind farm area. Annual status report 2003*. Report commissioned by Elsam Engineering A/S 2003. National Environmental. Research Institute, Rønde, Denmark.

Ramsar Convention Bureau (1988) Convention of Wetlands of International Importance especially as waterfowl habitat. Proceedings of the third meeting of the Conference of the Contracting Parties. Regina, Saskatchewan, Canada; 27 May to 5 June 1987. Ramsar Convention Bureau, Switzerland.

Sims, M., Wanless, S., Harris, M.P., Mitchell, P.A. and Elston, D.A. (2006) Evaluating the power of monitoring plot designs for detecting long-term trends in the number of common guillemot. Journal of Applied Ecology **45**, 537-546.

Skov, H., Durinck, J., Leopold, M.F. and Tasker, M.L. (1995) Important Bird Areas for seabirds in the North Sea, including the Channel and Kattegat. BirdLife International,Cambridge.

SNH (2002) Methodology for assessing the effects of wind farms on ornithological interests: methodology for assessing the effects of wind farms on ornithological interests. SNH and BWEA Report.

Stewart-Oaten, A., Murdoch, W.M. and Parker, K.R. (1986) Environmental impact assessment: ‘pseudoreplication’ in time? *Ecology*, **67**, 929-940.

Underwood, A. (1994) On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications **4**, 3-15.

United Nations. (1982). United Nations Convention on the Law of the Sea. UN Doc. A/Conf. 62/122.

Wind Directions, Editorial (2003) Offshore future. Projects Stack Up in Europe’s Seas. Wind directions, September issue: 18-33. European Wind Energy Association.

Winkelman, J.E. (1992) De invloed van de Sep-proefwindcentrale te Oosterbierum (Fr.) op vogels. DLO-Instituut voor Bos- en Natuuronderzoek. RIN-rapport 92/2-5, 4 volumes, Arnheim, Netherlands.

Zar, J.H. (1988) Biostatistical Analysis, 4th Edition. Prentice-Hall, Englewood Cliffs, NJ.