

University of St Andrews



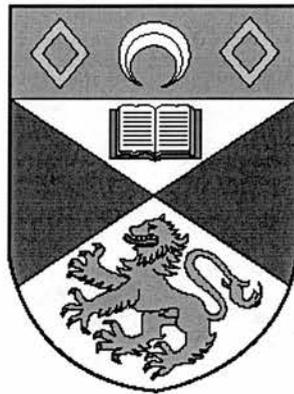
Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

Modelling heterogeneity in cetacean surveys

Sharon L. Hedley



Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

in the School of Mathematics and Statistics,

UNIVERSITY OF ST ANDREWS.

August, 2000.



Declarations

1. I, Sharon L. Hedley, hereby certify that this thesis, which is approximately 35,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date . 14/11/00 . signature of candidate

2. I was admitted as a research student in October 1995 and as a candidate for the degree of PhD in October 1996; the higher study for which this is a record was carried out in the University of St Andrews between 1995 and 2000.

date . 14/11/00 . signature of candidate .

3. I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date signature of supervisor

4. In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

date . 14/11/00 . signature of candidate .

Abstract

Methods for improving estimation of cetacean abundance from line transect and mark-recapture surveys are proposed.

Using either generalized linear or generalized additive models (GLMs or GAMs), two approaches are suggested which allow heterogeneity in the spatial distribution of cetaceans to be modelled from standard line transect data. In the first approach, the transect lines are divided into smaller discrete units, and the expected number of detections in each unit is modelled using explanatory spatial covariates. In the second approach, the response is derived from the observed waiting times (or distances) between detections. Fitting this model within the usual GLM or GAM framework would require restrictive assumptions, therefore an iterative procedure is formulated which enables a realistic model to be fitted. Alternatively, it is shown how this approach can be framed as a point process model, and it is suggested how the likelihood for the observed along-trackline distances could be maximized. The methods are illustrated using line transect data from a survey of Antarctic minke whales. A surface representing the variation in density throughout the survey region is obtained, from which abundance may be estimated by numerical integration. It is also shown how the first approach can be adapted for cue count data, using an example from an aerial survey for minke whales off West Greenland.

Heterogeneity in the capture probabilities of individuals is a potential source of bias in mark-recapture estimates. A method is proposed which allows this heterogeneity to be modelled, and accounts for the past capture histories of the individuals. The model may be fitted using logistic regression with an offset variable which must be estimated. Results from fitting the logistic regression model to photo-identification data from a study of bottlenose dolphins in the Moray Firth, Scotland, are presented.

Acknowledgements

Firstly, I should like to thank my two supervisors, Steve Buckland and David Borchers, for their willingness to discuss the research, providing encouragement and motivation, and supporting my interest in gaining practical experience in cetacean survey methods.

I am grateful to the International Whaling Commission for financial support, and to members of its Scientific Committee for feedback on the spatial modelling methodology as it was being developed, particularly Mark Bravington and Justin Cooke.

The work in Chapter 6 was conducted in collaboration with Ben Wilson, who kindly provided the photo-identification data.

I should like to thank the following friends and colleagues who have contributed in different ways, but have nevertheless helped me complete this thesis: James Busby; Liz Clarke; Rachel Fewster; Peter Jupp; Fernanda Marques; Clare Parnell; Samantha Strindberg and Linda Walton.

Finally, I should like to thank Tom for his support during what was a rollercoaster ride.

Contents

1	Introduction	1
1	General background	1
2	Preliminaries	2
3	Thesis outline	6
2	Line Transect Cetacean Surveys: a Review of Methods	9
1	Introduction	9
2	Methodology for Two Platform Data	10
2.1	Introduction	10
2.2	BT survey mode	12
2.3	A Comparison of IO Methods	13
2.4	Summary	20
3	Mean School Size Estimation	21
3.1	Introduction	21
3.2	Comparison of Methodology	22
3.3	Summary	26

4	Estimation by IWC Small Management Area	27
4.1	A Summary of Available Methods	27
3	Incorporating Spatial Variation	29
1	Motivation	29
2	Introduction to Spatial Modelling in Wildlife Surveys	30
3	A spatial modelling formulation for strip transect data	32
3.1	Modelling group density	32
3.2	Modelling animal density	35
4	A spatial modelling formulation for line transect data	36
4.1	Modelling group density	36
4.2	Modelling animal density	40
5	Variance estimation	41
5.1	The problem of dependent observations	41
5.2	The non-parametric bootstrap	42
5.3	The jackknife	43
5.4	The parametric bootstrap	45
6	Applications of the methodology	47
6.1	Line transects: a shipboard survey of Antarctic minke whales	47
6.2	Cue counts: an aerial survey of minke whales off West Greenland	54
7	Discussion	60
4	An Alternative Perspective: Using Intervals Between Sightings	63

1	Introduction	63
2	Constant expected encounter rate model	64
2.1	Derivation	64
2.2	A spatial model based on waiting distances	65
3	Variable expected encounter rate model	66
4	Variance estimation	67
4.1	The jackknife	69
4.2	The parametric bootstrap	70
5	Application of the methodology	70
5.1	Antarctic minke whales	70
6	Further research	75
6.1	A convergence problem	75
6.2	Discussion	76
5	Point Process Modelling of Line Transect Data	79
1	Introduction	79
1.1	Some definitions and notation	80
1.2	Putting these definitions into a line transect context	81
2	Model Formulation	83
2.1	A model based on detection locations	83
2.2	A likelihood for waiting areas	85
2.3	A note on maximising the likelihood	88

2.4	Maximisation of the likelihood: a simplified case	88
3	Some comparisons with conventional line transect estimation	89
3.1	A pdf for detection locations	89
3.2	Equivalence of the point process likelihood derivation and the conventional line transect estimator	90
4	Discussion and further work	91
6	Modelling heterogeneity in mark-recapture analyses	93
1	Introduction	93
2	Modelling heterogeneity in capture probabilities	95
2.1	The Alho-Huggins approach	95
2.2	A general logistic regression approach for modelling capture probabilities	96
2.3	An algorithm for fitting the logistic regression model	99
3	Application of the methodology	100
3.1	A photo-identification study of bottlenose dolphins	100
4	Discussion	105
	Appendices	109
	A: Inclusion of past capture histories when there are two sampling occasions	109
	B: An S-PLUS program to obtain abundance estimates using the logistic regression methodology	110
	C: Maple commands for use in implementing the logistic regression methodology	113

7	Summary and Conclusions	115
1	Spatial modelling	116
2	Mark-recapture modelling	119
	Bibliography	120

Chapter 1

Introduction

1 General background

In order to manage whale stocks successfully, an assessment of their size and status is required. Although line transect methods have been used in the assessment of birds and terrestrial animals since the late 1940s, the particular difficulties in surveying animals which inhabit the oceans, spending most of their lives underwater, had, until relatively recently, largely precluded the use of these methods to assess whales; mark-recapture methods were much more prevalent for this purpose. The influential monograph by Burnham *et al.* (1980) presented an integrated framework for the design and analysis of line transect surveys, instigating a notable increase in methodological and applied research in this field. Much of the development of line transect methodology, particularly that which is most useful for the assessment of whale stocks, has since been conducted under the aegis of the Scientific Committee of the International Whaling Commission (IWC).

Although line transect methods provide an estimate of the density of whales in a given area, together with an indication of the precision of such an estimate, they can only relate density in a crude manner to the whales' environment. In principle, stratification of the survey area into blocks representing different habitat types is one possible way to relate cetacean distribution to habitat, but in practice these blocks are required to be quite large, in order that a sufficiently high number of sightings can be made in each block to facilitate calculation of stratified estimates. There is, however, an increasing awareness

that the effects of environmental change on cetacean populations should be given high priority by the scientific community, and that assessment methods should attempt to incorporate variables related to the marine environment which these populations inhabit.

Spatial models are widely used in wildlife surveys to explain species distribution using habitat-related variables, particularly for terrestrial surveys where detailed environmental maps of the survey area may be provided by geographic information systems. The data from these surveys are typically available at gridded sample points throughout the area. It would be neither practical nor efficient to attempt to collect gridded data on a cetacean survey. As outlined above, line transect surveys have become an accepted survey technique in the assessment of cetaceans; it would therefore be desirable to develop methodology which integrates conventional line transect data, environmental (habitat-related) data, and data indicating prey availability, in order to model spatial variation in cetacean distribution and density.

The use of spatial models in the assessment of cetacean populations has other potential advantages. In the context of stock management, the IWC defines ‘Small Management Areas’ (SMAs) for which an estimate of the stock may be required. The boundaries of the SMAs are not necessarily coincident with the boundaries of the survey strata, and so any estimation method must provide a means to estimate abundance within a subset of a stratum, or across subsets of several strata. The use of conventional stratified line transect methods in this regard suffers from the same sample size restrictions that limit their suitability for use in relating whale density to environmental variables. However, spatial models enable density to be estimated across any subregion of interest. Furthermore, data from outside that subregion may be included in the spatial models; this should ensure a better fit and consequently more reliable inference. By incorporating the spatial component of habitat variability in the model, improved precision of the abundance estimates might also be expected.

2 Preliminaries

In this section, we outline some general introductory material on line transect methods, including some notation and fundamental concepts. It may also be found in several standard texts, such as Seber (1982) and Buckland *et al.* (1993a); these may be referred to

for more details.

Within some survey region, a line transect cetacean survey typically involves observers on a survey platform (such as a vessel, an aircraft or both), which travels along predetermined tracklines, laid out according to some appropriate sampling design. The observers record all whales that are seen. Modelling is based on the perpendicular distance of the whale pods from the trackline (Figure 1). Whilst it is not necessary to assume that the pods are distributed randomly throughout the survey region, the trackline design should be such that it may be reasonably assumed that the distribution of the perpendicular distances is uniform. This is an important assumption, since it effectively allows us to estimate the proportion of whales missed by the observers, and thus an estimate of all whales in the survey area can be obtained.

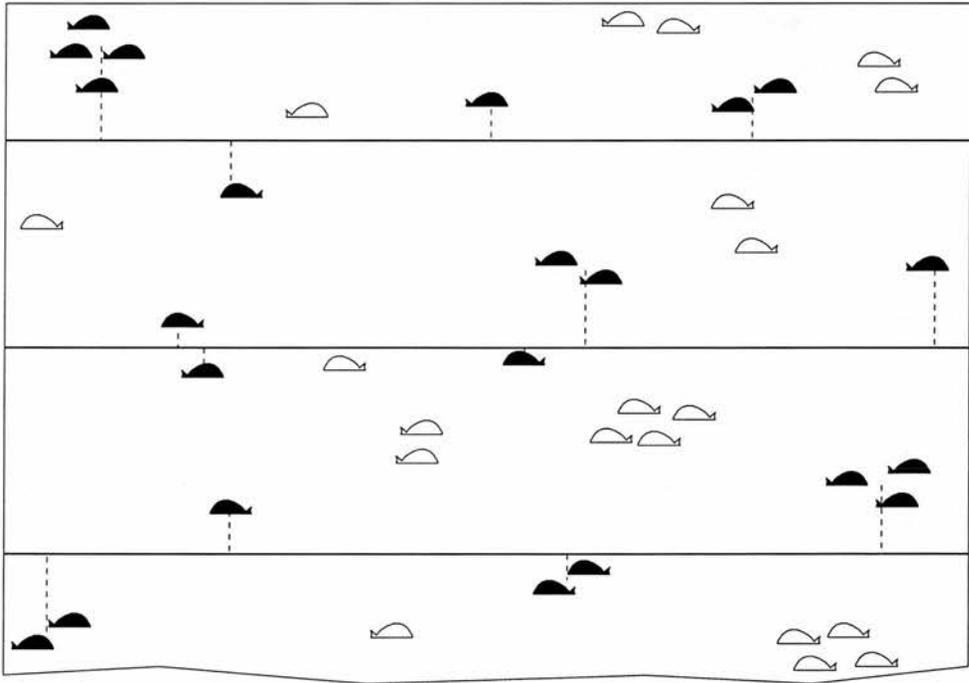


Figure 1: A typical survey area, covered by three horizontal transects represented by the thick solid lines, depicting detected (black) and undetected (white) whales. Line transect estimation is based on modelling the perpendicular distance of detected whales (or whale pods) from the transect line, shown here as dashed lines. Note that for a pod of whales, the perpendicular distance is that from the centre of the pod to the transect line.

Suppose for now that the observers are able to detect all whale pods out to some perpendicular distance, ω , either side of the trackline, and that the total length of trackline

surveyed is L . Then the area surveyed is a rectangular strip of area $2\omega L$ and the estimated density of pods, \hat{D} , within the strip is given by

$$\hat{D} = \frac{n}{2\omega L},$$

where n is the number of pods detected. The estimated density of whales \hat{D}_W is given by $\hat{D} \cdot \hat{E}(s)$, where $\hat{E}(s)$ is the estimated mean pod size. Given that the survey region is of total area A , then the estimated number of whales, \hat{N}_W in the region is

$$\hat{N}_W = \frac{n \hat{E}(s) A}{2\omega L}.$$

Since we have assumed that within the strip of area $2\omega L$ all pods were detected, this formula actually gives us the estimated number of whales from a *strip* transect survey. This survey method is rarely, if ever, used for surveying whales, since detections beyond ω are excluded from the analysis, and given the low densities at which whales generally occur, the method would be extremely inefficient. However, the concept of defining a strip within which all pods are detected is a useful one for explaining how density is estimated from a line transect survey.

Line transect methodology relies on estimating a detection function g , which (conventionally at least) is a function of the perpendicular distances, y_i , $i = 1, \dots, n$, such that

$g(y)$ = the probability of detecting a whale pod, given that it is at perpendicular distance y from the line.

In order to estimate this function, a probability density function (pdf), $f(y)$, is fitted to the frequency distribution of perpendicular distances. This pdf may be fitted to the distances from all n detections, but it is usually found that more robust estimation is obtained by truncating a small percentage of the sightings beyond some perpendicular distance, w (to reduce the influence of the tail on the overall fit). Assuming that pods on the transect line (ie. those at zero perpendicular distance) are detected with certainty, it can be shown that the detection function $g(y)$ is identical in shape to the pdf $f(y)$, but rescaled so that its maximum value (at zero distance) is 1 (see Figure 2).

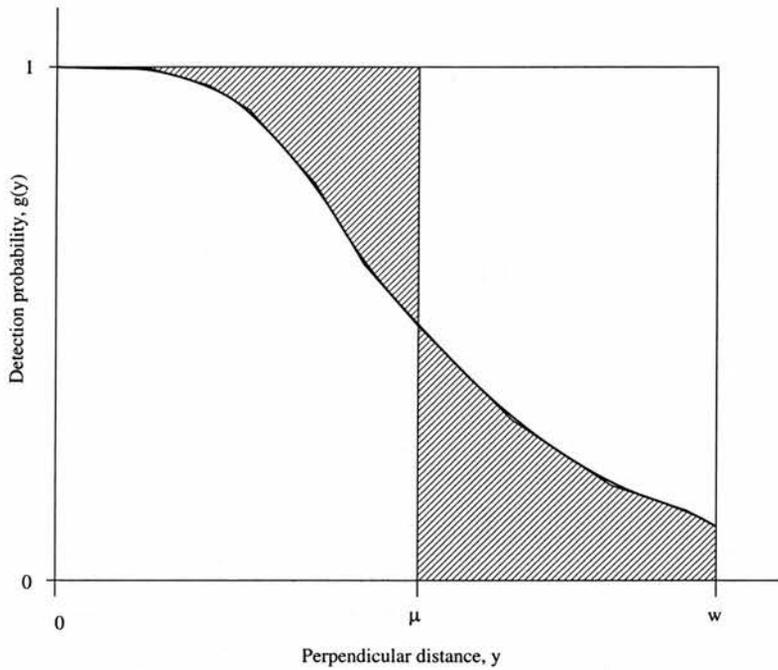
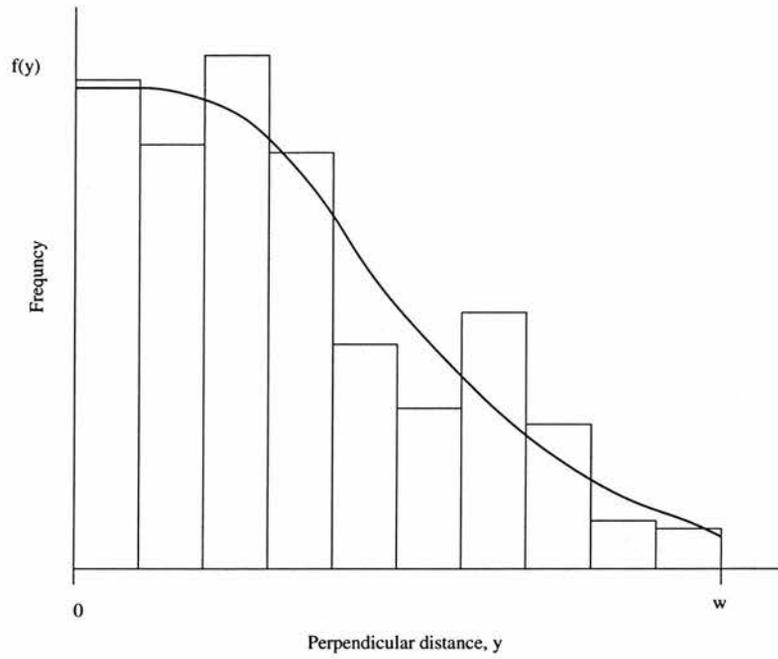


Figure 2: The top figure shows a pdf, $f(y)$, fitted to perpendicular distances, which have been truncated at distance w , and grouped into intervals for displaying as a histogram. The bottom figure shows the corresponding fitted detection curve, $g(y)$, with $g(0) = 1$. The shaded regions are equal in area; μ is thus the effective strip (half-) width.

Since $f(y)$ is a pdf, its integral, $\int_0^w f(y) dy$ equals 1. Since the shaded regions shown in the bottom half of Figure 2 are equal, then $\int_0^w g(y) dy = 1 \cdot \mu$. Therefore the relationship between the pdf of perpendicular distances and the detection function is given by

$$f(y) = \frac{g(y)}{\mu},$$

and in particular,

$$f(0) = \frac{1}{\mu}.$$

The quantity μ is termed the effective strip (half-) width. The expected number of whale pods missed when they are located at perpendicular distances less than μ is equal to the expected number that are seen at distances beyond μ .

Given a total transect length, L , the surveyed strip in a line transect survey is of estimated area $2\hat{\mu}L$, with $\hat{\mu}$ being estimated from the fitted pdf as $[\hat{f}(0)]^{-1}$, and the estimated abundance of whales in a survey region of area, A , is

$$\hat{N}_W = \frac{n \hat{E}(s) A}{2\hat{\mu}L}.$$

3 Thesis outline

One of the basic assumptions conventionally required for estimation by line transect methods is that objects (here, whale pods) on the trackline are detected with certainty. For many whale species, this assumption does not hold. Survey techniques in which sightings are made from more than one observation platform, and resultant analytical methods, have been developed to enable unbiased estimation. These are reviewed in chapter 2, along with methodology for estimating mean school size. The latter can be problematic because of the so-called ‘size bias’ effect, in which large schools tend to be detected at greater perpendicular distances from the transect line than small schools, on average. Much of the content of this chapter is drawn from Hedley *et al.* (1997a).

In chapter 3, an approach for fitting spatial models to line transect data within a generalized linear or a generalized additive modelling (GLM or GAM) framework is proposed.

The models are based on observations obtained from dividing the transects into smaller units, called ‘segments’, and modelling the number of detections in each segment. The methodology is applied to minke whale sightings data from a shipboard line transect survey in the Southern Ocean, and to cue count data from an aerial survey of minke whales off West Greenland.

In chapters 4 and 5, an alternative spatial modelling approach is suggested. In both cases, the response variable for the spatial model is based on the along-transect distance between successive detections. Chapter 4 primarily deals with applying the approach using standard software, and as in chapter 3, methodology is proposed which enables the models to be fitted within either a GLM or a GAM framework. An iterative procedure is proposed which allows between-detection variability in density, encounter rate and explanatory spatial variables to be incorporated. The methodology is applied to the Southern Ocean minke whale data and the results are compared with those from the previous chapter. The method is unsuitable for cue count data. Parts of chapters 3 and 4 form the basis of Hedley *et al.* (1999); the West Greenland minke whale analysis was presented in Hedley *et al.* (1997b).

Chapter 5 covers the underlying theory behind the approach of chapter 4, developing a likelihood framework based on point process models. The material in this chapter differs somewhat from the previous two, being largely theoretical in nature. However, like the others, it was developed with real-life applications in mind, and thus builds upon previous attempts to frame line transect estimation within point process modelling (Stoyan, 1982; Högmänder, 1991; Högmänder, 1995), in which the simplifying assumption of a homogeneous distribution of points (here, whale pods) was made.

In contrast to all previous chapters, the methodology proposed in chapter 6 is for mark-recapture rather than line transect data. This stand-alone chapter was motivated by the problem of modelling heterogeneity in the capture probabilities of individual bottlenose dolphins in a photo-identification study, in particular, when one of the aims of the study was to examine the dynamics of the population and thus, it was thought more appropriate to use open, rather than closed, population models. Since such models are unable to incorporate heterogeneity in individuals’ capture probabilities, methodology based on previous work by Huggins (1989; 1991), Alho (1990) and Buckland *et al.* (1993b) is developed, which allows closed population models which incorporate this heterogeneity to be fitted

using standard statistical software. It is suggested that these models are combined with the 'robust design' of Pollock (1982).

Finally, in chapter 7, unresolved issues are discussed and suggestions for future research are presented.

Chapter 2

Line Transect Cetacean Surveys: a Review of Methods

1 Introduction

In order to apply the basic line transect methods outlined in Chapter 1 to estimate cetacean abundance, the following assumptions must be met (Buckland *et al.*, 1993a):

1. Animals on the trackline are detected with certainty.
2. Animals are detected at their initial location.
3. Measurements are exact.

Since Assumption 1 is clearly invalid for many cetacean species, methods have been developed to try to estimate the probability of detection on the trackline. These methods are reviewed in the next section. Violation of Assumption 2 is not problematic provided that the direction of animal movement prior to detection occurs independently of the observer, and at a speed that is slow relative to the speed of the observation platform. However, if movement is in reaction to the observer, density estimation is biased. A survey method that can account for such movement is described in section 2.2. Failure of Assumption 3 is not discussed in detail here since field procedures which ensure the accuracy of sighting distances and angles should always be implemented. However if this is not possible,

methods do exist to attempt to accommodate poor quality measurements. These include *ad hoc* methods such as ‘smearing’ (Butterworth, 1982; Buckland and Anganuzzi, 1988), and a recent theoretically based approach by Chen (1998) which assumes an exponential power series detection function, and corrects for bias in the estimation using method of moment estimators.

Since many cetacean species occur in pods or schools, the sighting unit is usually the school and individual abundance can be estimated by the product of the school density estimate and an estimate of mean school size. On average, relatively large schools may be detected at greater perpendicular distances than small schools, therefore using the sample mean of detected schools as an estimate of mean school size is often unsatisfactory. In section 3, methods which address the problem of estimating mean school size when detectability is a function of school size are presented.

Standard line transect methods yield a single estimate of average abundance over some pre-defined survey stratum, but say nothing about density might vary within that stratum. In some circumstances it may be desirable to have estimates of abundance for areas which do not necessarily correspond, or even share common boundaries, with these strata. For example, implementation of the Revised Management Procedure (RMP) of the International Whaling Commission (IWC) requires estimates to be available by small geographic regions termed ‘Small Management Areas’ (SMAs) which may be subregions of a survey stratum or may intersect several strata. Three possible approaches to this problem are briefly outlined in section 4.

2 Methodology for Two Platform Data

2.1 Introduction

Initial attempts to estimate the probability of detection on the trackline, $g(0)$, were based on the models of Doi (1971; 1974). These were exceptionally detailed models, giving rise to estimates of $g(0)$ with very high estimated precision, but the robustness of the approach was shown to be unsatisfactory (Best and Butterworth, 1980; Doi *et al.*, 1982; Doi *et al.*, 1983). In an attempt to achieve more robust estimation when detection on the trackline is uncertain, the two main approaches that have been taken are ‘variable effort’ (VE)

methods and ‘independent observer’ (IO) methods.

VE methods estimate $g(0)$ using data from surveys of the same region at different levels of effort. In practice this has usually meant using data from surveys at two different platform speeds, although Zahl (1989) proposed varying the number of observers. Polacheck (1995) reports the results of an experiment carried out to investigate the effect on $g(0)$ of varying the number of observers (specifically by increasing search effort on the trackline — itself not without problems, see Buckland *et al.* (1993a, p30-31)). Polacheck concludes that because of individual observer heterogeneity, in general the variation in sighting efficiency caused by the different number of observers cannot be quantified and is thus an inappropriate method by which to estimate $g(0)$. Butterworth *et al.* (1982), Cooke (1985) and Zahl (1989) develop variable speed estimators of $g(0)$. The estimators were found to perform poorly, yielding 95% confidence intervals which were sometimes wider than the feasible range for $g(0)$. Serious failures in the assumptions of the methods were identified and the methods were abandoned in favour of IO methods.

IO methods use data from simultaneous surveying by two observation platforms. The platforms may be located on the same vessel, on two separate vessels, on a vessel and a helicopter or fixed-wing aircraft, or on two separate aircraft. Historically, the methods required each platform to operate independently of the other, but in recent years methods have been developed which require only one-way independence (see section 2.2). The methods are analogous to two-sample mark-recapture methods where sighting by one observation platform constitutes ‘marking’, and sighting of the same animal by an observer on the other platform constitutes ‘recapture’. In a line transect context, ‘duplicates’ correspond to recaptures. (A duplicate sighting is one which has been seen independently by both platforms.) IO methods currently fall into one of two categories. In the first, the target animals are assumed to be continuously available for detection, while in the second the animals are assumed to be available at discrete points in space and time.

Methods which have been proposed for estimating abundance and/or $g(0)$ from IO data assuming continuous availability include the ‘Product’ method (Butterworth *et al.*, 1984; Kishino *et al.*, 1986; Buckland, 1987), the ‘Modified Product’ method (Palka, 1993, called the ‘Direct Duplicate Method’ by Palka), the ‘Direct’ method (Butterworth and Borchers, 1988; Hiby and Hammond, 1989), the ‘Buckland-Turnock’ method (Buckland and Turnock, 1992) and the ‘Modified Binary Regression’ method (Borchers *et al.*, 1998).

Two other methods for the analysis of IO data are mentioned here for completeness. The first, proposed by Butterworth (1991), suggests \hat{D}_I/\hat{D}_{all} as a positively biased estimator of $g(0)$ for platform 1, where \hat{D}_I is the conventional line transect estimator of density from data from observation platform 1 only, and \hat{D}_{all} is the conventional line transect estimator of density from all platforms. The second is a removal method based on the assumption that $g(0)$ from both platforms is equal (Barlow, 1993).

Initial IO analysis methods relied on the assumption of unconditional independence of detections between platforms. This assumption is violated if the probability of detection depends on variables not incorporated in the detection function model. In contrast to conventional line transect estimators (with $g(0)=1$), IO estimators are not pooling robust and unbiased estimation from IO survey data requires that all variables which substantially affect detection probability are incorporated in the estimation of detection probabilities.

With discrete animal availability, detection probability is dependent upon the pattern of availability of the animals. Schweder (1974; 1977) proposes models for IO discrete availability data which, unlike continuous availability models, do not require detections of animals to be independent between platforms. Instead, the models assume that sightings of cues by each platform are independent. These early models did not provide estimates of abundance or $g(0)$, but in later papers, such models are developed and successfully implemented (see Schweder (1990) for theory, and Schweder *et al.* (1991) for an application). Cue-based models provide estimates of cue density, which must be combined with an estimate of cue rate to obtain an estimate of animal density. Reliable estimation of cue rate is difficult and its incorporation into whale abundance estimation requires additional assumptions. Usually cue rate is estimated by direct observations, or satellite-tagging, of a small number of whales for a period of time. One assumption is therefore that the animals on which the estimate of cue rate is based exhibit surfacing and diving behaviour typical of the whole population under study.

2.2 BT survey mode

In standard IO survey mode, the observers on the two platforms search the same area independently of each other. An alternative two-platform survey mode which only requires one-way independence between the platforms was proposed by Buckland and Turnock

(1992). In this survey mode, observers on one platform (the ‘secondary’ or ‘tracking’ platform), search an area ahead of the area where animals may be detected by observers on the other platform (the ‘primary’ platform). The area searched by observers on the tracking platform must be sufficiently wide to ensure that schools which move into the area for potential detection by the primary observers have an opportunity to be tracked. The primary platform observers search independently of those on the tracking platform. Buckland and Turnock’s original application of this survey mode (termed BT mode) is in an analysis of Dall’s porpoise (*Phocoenoides dalli*) data where the primary platform is on board a ship, and the tracking platform is on a helicopter flying slightly ahead of the ship. However, BT mode may also be conducted on board a single vessel, for example if observers on the tracking platform search with binoculars and those on the primary platform search by naked eye (see for example, Hammond *et al.* (1995)). The main reasons for employing BT survey mode are as follows:

(i) With suitable analysis methods, BT survey mode is able to yield asymptotically unbiased abundance estimates in the presence of responsive and/or random animal movement. The animals must, however, be detected by the tracking platform prior to any reactive response.

(ii) Ensuring that the two platforms do not search the same area of sea at the same time helps reduce the dependence between detections which can arise from unmodelled variables such as animal surfacing behaviour.

2.3 A Comparison of IO Methods

The problem of uncertain duplicate identification affects all two-platform estimators and is worse in high density areas. Improvement in practical survey techniques can help, but there is a need for models and estimation methods which incorporate this uncertainty in a statistically rigorous manner. Substantial progress in addressing this issue has been made by Hiby and Lovell (1998) and by the IWC Abundance Estimation Working Group (Cooke, 1997; IWC, 1997; Schweder *et al.*, 1997). Hiby and Lovell (1998) develop an innovative method for aerial surveys which does not require identification of individual duplicate pairs (see section 2.3.5), whereas IWC (1997) concentrates on explicit and objective rules for identifying duplicate pairs based on the relative proximity in time and

space of surfacings. Schweder *et al.* (1997) provide details of two possible algorithms for automated identification of duplicates. Both of these attempt to identify duplicate ‘tracks’ made by the same whale and seen from both platforms. A different approach is taken by Cooke (1997). His method relies on automated identification of duplicate cues rather than duplicate whales — a substantially easier task provided that sufficient accuracy in the data (and particularly in recorded sighting times) can be obtained. However, all automated methods developed to date have been specifically designed to identify duplicates when the field procedures are such as those conducted on surveys of North Atlantic minke whales. The general applicability of these methods to data from other surveys is questionable, in particular for species or stocks that do not usually occur singly or in small groups. Errors in estimating radial distances and angles to sightings, and in recording sighting times, clearly play an important role in the successful assessment of duplicate status. When duplicates are classified by an automated algorithm, measurement errors become even more critical. In an attempt to reduce the effect of such errors, measurement error models have been devised and implemented (Schweder (1996; 1997)). These models rely on data collected from buoy sighting experiments, similar to the Estimated Angle and Distance Experiments carried out on, for example, IWC/IDCR (IWC/International Decade of Cetacean Research) surveys. The practical constraints of any such experiment make simulation of true survey conditions impossible — the buoy is an unrealistic representation of a whale cue and experimental conditions are often very different from survey conditions. The difficulties in obtaining accurate estimates of angle and distance have further implications in the way they affect the form of the detection function. There is a need for further research in this area.

Leaving uncertainty in duplicate identification aside, an assessment of two-platform methods should consider the generality of the assumptions on which the methods are based and their robustness to violation of these assumptions. All double platform estimation methods are based on assumptions of independence between detections by the two platforms, and all such estimators are biased by failure of these assumptions if $g(0) < 1$. The degree to which the independence assumptions implicit in the various methods are likely to be violated is therefore a central criterion on which to evaluate the methods. Another important criterion is the degree to which the methods are able to accommodate random or responsive animal movement; the latter can be a major source of bias (Buckland and Turnock, 1992; Buckland *et al.*, 1993a), and when $g(0)$ is less than unity, the former is

inextricably linked with its estimation. This is because the perpendicular distances of sightings may change appreciably between the time they are detected by one platform and the time they are detectable or most detectable by the other platform (see Buckland and Turnock (1992) and Buckland *et al.* (1993a) for a fuller discussion of this point). It has long been recognized that $g(0)$ estimates are biased when there is unmodelled heterogeneity (see for example, Butterworth and Borchers (1988)), therefore a third criterion when assessing two-platform methods is the means by which they are able to incorporate additional explanatory variables (covariates) to model any heterogeneity not accounted for in the perpendicular distance model of the detection function. The methods are discussed in a little more detail below, noting their primary strengths and failings.

2.3.1 Product Methods

The original Product method (Butterworth *et al.*, 1984; Kishino *et al.*, 1986; Buckland, 1987) relies on a strong assumption of unconditional independence between detections. The probability of both platforms detecting an animal at a given perpendicular distance is assumed to be equal to the product of the detection probability functions for each platform. Analyses of IWC/IDCR IO data using this method were suspended largely because the strong independence assumptions did not appear to be met. Palka's (1993) Modified Product method is based on a weaker assumption of unconditional independence, namely that detections by different platforms of animals on the trackline (but not necessarily elsewhere) are unconditionally independent. The Modified Product method is able to relax the independence assumption of the Product method by estimating the shape of the duplicate detection function from the perpendicular distance distribution of observed duplicates, rather than assuming that it is determined by the product of the two platforms' detection functions. (The duplicate detection function is the probability of both observers detecting a school, as a function of perpendicular distance.) In some applications the assumption of unconditional independence on the trackline may be violated. For example, larger schools on the trackline will generally be more detectable than smaller schools. The extent to which Modified Product Method estimates are biased by violation of this assumption remains to be investigated. A practical limitation of the method is that it will not always be logistically possible to obtain sufficient duplicate detections for reliable estimation of the shape of the duplicate detection function. Recently, Laake (?) developed a likelihood framework for the Modified Product method, removing this limitation.

The BT estimation method was developed specifically to accommodate responsive animal movement in estimation. In BT survey mode, the secondary platform may be thought of as confirming the position of a sample of animals before they react, and the proportions of these detected by the primary platform allow estimation of the primary platform detection function, with $g(0) \leq 1$. If detections from the secondary platform are made prior to responsive animal movement, the method is asymptotically unbiased. When animals are stationary and there is no unmodelled heterogeneity, the BT analysis method is similar to the Modified Product method, in that both methods use the distribution of observed duplicates to estimate the shape of the duplicate detection function. The essential difference between them in this case lies in the way that the perpendicular distance density of observed primary platform detections is estimated. In practical applications, the fact that estimation using the BT method assumes that the probability, g_{PS} , of detecting an animal from both platforms is equal to the product of detecting the animal individually from the two platforms ($g_P \cdot g_S$) for *all perpendicular distances* often leads to negatively biased estimation. In contrast, the Modified Product method only requires this assumption to hold on the trackline, i.e. at zero perpendicular distance.

Assuming that heterogeneity can be modelled, for example using covariates, a merit of the BT method is that it allows for the fact that animals may have moved from the position at which the secondary platform saw them by the time they are visible to the primary platform. In contrast, neither the Product method nor the Modified Product method are able to accommodate random or responsive animal movement.

2.3.2 Binary Regression Methods for Ungrouped Data

The Modified Binary Regression method for estimating detection functions is essentially a binary regression method which allows conventional generalized linear model (GLM) link functions for binary data to be modified to obtain improved models for line transect detection functions. If no such modification is necessary, the method reduces to using standard binary regression to estimate the detection function. The method was designed for use with BT survey mode to allow for animal movement and it is therefore not symmetric across platforms. Data from the tracking platform are used only to generate trials for the primary platform. A sighting first detected by the tracking platform sets up a Bernoulli trial. Each trial has a vector of explanatory variables associated with it, together with a

binary response (1 if the sighting was detected by the primary platform and 0 if it was missed). The detection function for the primary platform can be estimated from these data using one of the many statistical software packages that is capable of generalized linear modelling. If standard GLM link functions are inadequate, the link function may require modification. The method described in Borchers *et al.* (1998) is one way of modifying standard link functions.

The method can readily be modified to deal with the symmetric case in which, as in standard IO analyses, detection functions for both platforms are estimated, although in this case the method cannot accommodate animal movement. With IO mode data, the estimation problem is no longer a straightforward binary regression problem, even when standard link functions are used. However, standard statistical software can still be used in an iterative manner to estimate the detection functions for both platforms. Borchers *et al.* (1998) describe the method, which is an adaptation of the methods of Huggins (1989), Alho (1990) and Buckland *et al.* (1993b) to the line transect context.

Once the detection function(s) have been estimated, an estimate of the detection probability for each detected school is calculated, conditional on the explanatory variables associated with the school. Abundance is estimated using an estimator similar to the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), in which these estimated detection probabilities are the ‘inclusion probabilities’.

2.3.3 The Direct Method for Grouped Data, and a Generalized Direct Method

Consider a case in which detection probabilities depend only on perpendicular distance. BT survey mode is used so that the tracking platform serves only to generate trials for the primary platform, and detections are grouped into perpendicular distance ‘bins’. Given only that a school falls in a particular bin, detections by the two platforms will not in general be independent, because the bin contains a range of perpendicular distances and therefore detection probabilities will change within each bin. A special case in which unconditional independence within a bin does hold is when detection functions are constant within the bins. The less the detection functions change over the range of a bin, the more detections by different platforms approach independence. The Direct method of Butterworth and Borchers (1988) is an approach which assumes unconditional independence of detections within bins. By making the bins sufficiently narrow, bias caused by different

detection probabilities within each bin should be negligible. In the limit, as bin width approaches zero, this binned estimation method approaches a binary regression method.

If the detection functions depend on variables other than perpendicular distance, this too will cause the estimates to be biased, unless this dependence is modelled. One way in which the Direct method can be generalized to incorporate additional explanatory variables is to group the data into bins with respect to both perpendicular distance and the other explanatory variables. Independence of detections within each bin is then obtained if the detection functions are constant with respect to both perpendicular distance and the explanatory variables within each bin. A convenient way of estimating the detection function from such binned data would be to treat the midpoint value (say) of explanatory variables in each bin as the explanatory variable, and use GLM estimation software for binomial response models. Like the binary regression methods, this sort of generalized Direct method could be used in conjunction with BT mode survey to accommodate responsive or random whale movement as well as additional explanatory variables. However, whilst such an approach may be convenient, it will at best be an approximation to the real situation because of the lack of independence within bins. The binary regression methods are preferred because they do not suffer from this problem.

2.3.4 Discrete Availability Methods

A method which accounts for the discrete availability of cetaceans was formulated by Schweder (1977; 1990) for IO data. It is most appropriate for analysis of data where the species in question occurs singly or in small groups which exhibit synchronous surfacing and diving behaviour. Such species give discrete cues. Species that occur in larger groups or do not dive synchronously tend to provide more, and often continuous, cues to the observers. The Discrete Availability (DA) method may be thought of as a two-stage process. In the first stage, the instantaneous detection function of cues, based upon their positions relative to the observers, is estimated. The second stage involves estimation of the detection function by integration over the along-trackline direction. The latter requires information on the surfacing and diving behaviour of the animals.

Given that the data are collected in IO mode and that the time and position of all cues (sightings *and* resightings) are recorded as they are seen by observers on each platform, then from the perspective of each platform, the positions of the cues relative to the other

platform can be calculated. The relative positions of cues seen by observers on one platform but either not seen by observers on the other platform, or only seen by observers on the other platform for the first time, are used to estimate the conditional sighting probability, $Q(x, y)$, of a cue at location (x, y) , given that it had not been previously observed. (The coordinate system is set up so that the origin is placed at the vessel, the positive x-axis is along the trackline in the direction of the vessel, and the y-axis is orthogonal (Schweder, 1990)). Each cue may have additional covariate information associated with it, characterizing its objective visibility, and this information can be incorporated into the estimation process, as in Cooke (1996) for example.

In order to obtain an estimate of the detection function, Schweder (1990) implements a computationally intensive Monte Carlo procedure using dive cycle patterns of tagged North Atlantic minke whales with $Q(x, y)$ estimated from a Bernoulli likelihood. The estimation method has been refined considerably since then, and now incorporates components relating to uncertainty in estimating the position of animals. As mentioned in section 2.3, automated duplicate identification algorithms have been developed for DA methods. By dispensing with the need for subjective judgements on duplicate status, the analysis can also take account of the probability of duplicate misclassifications. Cooke (1997) and Schweder *et al.* (1997) provide recent descriptions of the method.

Skaug (1997) and Skaug and Schweder (1999) develop a likelihood-based method with whales rather than cues as the detection unit which incorporates a model for whale surfacing behaviour. Duplicate whales are identified using an automated algorithm, as described in Schweder *et al.* (1997). The method is much less computationally intensive than the DA method for cues, but remains most appropriate for the analysis of data where the species in question occur primarily as single animals, such as North Atlantic minke whales.

2.3.5 IO Methods for Aerial Surveys

In their analysis of the SCANS (Small Cetacean Abundance in the North Sea) aerial survey data, Hiby and Lovell (1998) present a method that attempts to incorporate uncertainty in duplicate identification into the abundance estimation procedure. In the SCANS survey, the survey platforms were on board two fixed-wing aircraft which flew almost identical tracklines, one behind the other, approximately 9 km apart. The method does not require duplicate identification during a survey — in fact, the method does not even require dupli-

cates to be identified in the conventional sense at all. Subjective judgements are dispensed with and instead duplicates are ‘identified’ within a stochastic modelling framework, where the probability that two given sightings are assigned to be duplicates depends on their respective locations, the time interval between them, an assumed (gamma) distribution of swimming speeds of the animals in question and the density of animals in the area where the sightings occurred. Other variables, such as diving behaviour, can readily be incorporated into the model. For further details of the practical survey methods and the methodology used in their analysis, see Hiby and Lovell (1998) or Hammond *et al.* (1995).

2.4 Summary

All product methods rely (to varying degrees) on the relatively strong assumption of unconditional independence between platforms. The Product method itself assumes that this holds for all detections. Because of unmodelled heterogeneity, this assumption is likely to be invalid in practice, leading to biased estimation. With the Modified Product method, the assumption is relaxed so that the unconditional independence must only hold for detections on the trackline. This may yield biased estimators if there is unmodelled heterogeneity on the trackline. The BT method is also unable to model heterogeneity. Amongst the product methods, only the BT method is able to accommodate animal movement. Estimates from the other two methods may be substantially biased in the presence of responsive or random animal movement.

The Modified Binary Regression method supersedes the BT method and other product methods as a generally applicable method in that it is able to accommodate both animal movement and heterogeneity due to variables other than perpendicular distance. The Direct method can be generalized to accommodate both animal movement and heterogeneity due to variables other than perpendicular distance, but it relies on an assumption of conditional independence within bins which will often be violated.

The cue-based discrete availability models of Schweder *et al.* (1997, and earlier papers) are extremely computationally intensive and the high level of numerical complexity makes any assessment of the reliability of the procedures difficult. The whale-based models developed by Skaug (1997) and Skaug and Schweder (1999) provide a simpler framework for modelling discrete availability IO data at the expense of a few more assumptions. However, neither

of the two methods is able to accommodate responsive movement and in addition they both assume that there is no reaction to the observer in terms of surfacing behaviour. Since North Atlantic minke whales tend to occur as single animals and provide discrete cues, the models have proven to be effective for estimation of the abundance of this stock, but as a general method for many other species, the Discrete Availability method is likely to be inappropriate.

The IO aerial survey method has the advantage over shipboard surveys that provided the two aircraft are sufficiently separated, responsive movement is likely to be far less problematic. (In any case, whales may respond less to aircraft than to ships.) The stochastic framework avoids the need to identify duplicate pairs. The approach is relatively new, but the results from the SCANS survey (Hammond *et al.*, 1995) appear promising.

3 Mean School Size Estimation

3.1 Introduction

Modes of survey during a cetacean line transect fall into two general categories: Closing mode and Passing mode (of which IO mode and BT mode are special cases). In Closing mode, the observation platform (say, vessel) diverts from its course when a sighting is made, to confirm species identification and school size. In Passing mode, the vessel continues on its original course at all times, making species identification more difficult and often resulting in under-recording of the true school sizes. Failure to recognize this problem in estimation of mean school size would result in negatively biased estimates of abundance. Furthermore, estimation of mean school size should take into account the long-acknowledged problem of ‘size bias’, namely that the mean size of detected schools in a survey is a positively biased estimate of mean school size in the population when larger schools are more detectable than smaller schools. Up until 1988, analyses of minke whale data from IWC/IDCR surveys (which operate approximately half the time in Closing mode and half the time in Passing mode) estimated mean school size as the average size of schools detected within a perpendicular distance of 1.0 nautical miles from the trackline, using Closing mode sightings only. This approach is called the truncation method and is considered again briefly in section 3.2.1. Butterworth (1988) notes that the method

may be positively biased for estimating mean school size for minke whales from these surveys and proposes an *ad hoc* estimator of school size which involves fitting a detection function to the observed perpendicular distances of individual whales as well as to schools. This method was adopted by the IWC Scientific Committee until 1997, when, following a review of available methods (Borchers, 1994), a method based on a regression of $\log(\text{school size})$ against estimated detection probability, $\hat{g}(y)$ (Buckland *et al.*, 1993a), (where y is perpendicular distance from the trackline), was recommended.

Other methods to estimate mean school size are also considered in this section. Drummer and McDonald (1987) develop estimators of mean school size based on a bivariate detection function $g(y, s)$, where s is school size. Ramsey *et al.* (1987) propose a comparable method, not specifically to estimate mean school size in the population, but one in which mean school size, along with other covariates, could be included in the modelling process. Quang (1991) uses a nonparametric approach to model $g(y, s)$. More recently, Chen (1994) and Chen and Polacheck (1996), propose a kernel estimator based on explicitly modelling the probability of detecting a school as a function of both its perpendicular distance from the trackline and its size. Notable for its ability to estimate mean school size when detectability on the trackline is not certain (when two-platform data are available), the Horvitz-Thompson estimator (1998) is described in section 3.2.3.

3.2 Comparison of Methodology

3.2.1 Truncation, Stratification, Individual-based and Regression Methods

The simplest method for estimating mean school size in the presence of size bias exploits the fact that size bias in detected schools does not occur at distances from the line for which detection is certain. Thus expected school size, $E(s)$, may be estimated by the mean size of schools detected within a certain truncation distance, w' , where $g(w')$ is close to unity. Selection of the truncation distance is a balance between reducing size-biased selection of schools and maintaining a sufficiently large sample size for adequate precision. In situations where few school sizes are confirmed and/or the detection function falls off rapidly, this estimator is likely to perform poorly.

An alternative approach was considered by Quinn (1985) who evaluated estimators based on post-stratification by school size. The selected model is then fitted independently to

the data in each stratum. The choice of strata is determined largely by sample size. Each stratum should have a sufficiently large sample size for the effective search width to be reliably estimated without having such a wide range of school sizes within each stratum that intra-stratum size bias becomes an issue. A third method that has been considered bypasses the need to estimate mean school size at all. With this approach, estimation is implemented by applying line transect methods to individual whales rather than schools. Buckland *et al.* (1993a) state that this method is not to be recommended in general, although it may be useful if the population being sampled occurs in loose aggregations, rather than tight, easily defined schools. This method may perform poorly unless sample sizes are reasonably large. A fourth approach is to regress s or $\log(s)$ against y or estimated detection probability, $\hat{g}(y)$, with no (additional) truncation or stratification of the data.

Since the relationship between s and y may be substantially non-linear if the detection function has a shoulder, regression methods which assume a linear relationship between s and y are usually inappropriate. Buckland *et al.* (1993a) suggest that a regression of s against $\hat{g}(y)$ is more reasonable. Further, they note that the effect of increasing variance with increasing school size can be reduced by performing a regression of $\log(s)$ against $\hat{g}(y)$. All regression methods assume that the probability of detecting a school on the trackline is unity, regardless of its size. For many species this is an invalid assumption. Although precision is relatively good, variances are conditional on the estimated parameters of the detection function. Chen and Polacheck (1996) report that the coefficients of variation may often be underestimated by regression estimators because of violation of the distributional assumptions underlying the methods. However, this difficulty is readily solved by using a two-stage bootstrap algorithm, whereby $\hat{g}(y)$ is re-estimated prior to the regression in each bootstrap resample.

An appealing extension of the regression methods outlined above is the development of either a generalized linear or generalized additive modelling approach to estimate mean school size. Both of these provide greater flexibility and allow covariates other than perpendicular distance (or a function of perpendicular distance) to be included in the model. Research is ongoing in this area. Marques (in prep.) develops the approach, applying it to models of dolphin abundance in the eastern tropical Pacific.

3.2.2 Methods Based on Modelling the Bivariate Detection Probability, $g(y, s)$

The aim of these methods is to model directly the underlying distribution of sightings. Drummer and McDonald (1987) incorporate school size as a scale parameter in the detection function, which is assumed to be parametric in form. An analogous approach is developed by Ramsey *et al.* (1987), who note that the effective strip width ('effective area surveyed' in their terminology), α , where $\alpha = \int_0^{\infty} g(y)dy$, can be modelled as a function of covariates other than perpendicular distance, one of which may be school size, using multiple linear regression. Effective strip width is then incorporated as a scale parameter in the detection function, which is also assumed to take a parametric form. Parameters are estimated using the method of maximum likelihood. Beavers and Ramsey (1998) employ a different covariate adjustment approach from that of Ramsey *et al.* (1987), which can be implemented using Distance (Thomas *et al.*, 1998) together with any statistical software which performs multiple linear regression. Non-parametric forms for the detection function, such as Fourier series models, have also been proposed (Quang, 1991), but both the parametric and nonparametric forms assume that all schools on the trackline are detected, irrespective of their size.

3.2.3 The Horvitz-Thompson Method

If two-platform data are available, the Horvitz-Thompson method developed by Borchers *et al.* (1998) provides the best available method of estimating mean school size when detection on the trackline is uncertain. If only single platform data are available, then like the other methods, estimation assumes that animals are detected on the trackline with certainty. For the two platform case, mean school size is estimated by

$$\hat{E}[z_1] = \frac{\hat{N}_{indiv}}{\hat{N}}, \quad (1)$$

where the estimators of individual abundance, \hat{N}_{indiv} , and of school abundance, \hat{N} , are given by

$$\hat{N}_{indiv} = \sum_{i=1}^n \frac{z_{i1}}{\hat{P}(z_i|\hat{\theta})} \quad \text{and} \quad \hat{N} = \sum_{i=1}^n \frac{1}{\hat{P}(z_i|\hat{\theta})}. \quad (2)$$

where

z_{i1} is the size of the i^{th} detected school,

n is the total number of schools detected, and

$\hat{P}(z_i|\hat{\theta})$ is the estimated probability that at least one observer detects the i^{th} school, that school having a vector of covariates, z_i , associated with it. $\hat{\theta}$ represents the estimates of the parameters of the detection function.

Similar estimators are readily obtained for the single platform case. Variances may be calculated analytically or by bootstrapping, which though computationally intensive, has the advantage of not requiring an assumption of independence between the estimators of mean school size and school abundance. Borchers *et al.* (1998) recommend using transects as the sampling unit, with the caveat that if only a small number of transects is covered, the resulting bootstrap estimates may be unreliable. However in general, the bootstrap variance estimator is to be preferred. Indeed if it is used, and when observed school sizes must be estimated (as is often the case in practice), Borchers *et al.* (1998) note that the variance due to estimating a correction factor for bias in the observed school size estimates can be incorporated into the variance for estimated mean school size.

The Horvitz-Thompson estimator of mean school size is the only estimator which does not assume independence between school size and encounter rate. The use of this estimator in the analysis of the SCANS survey data (Hammond *et al.*, 1995; Borchers *et al.*, 1995) leads to what may appear to be a counter-intuitive result, namely that for certain species in some survey strata, abundance of individuals can be estimated more precisely than abundance of schools. One explanation for this effect is that in areas of high school density, animals may occur in smaller groups, and *vice versa*. This would occur if, for example, animal density is roughly constant, but in some regions or at some times, the animals aggregate into fewer, larger schools than in other regions or at other times.

3.2.4 The Kernel Estimation Method

The kernel estimation method (Chen, 1994; Chen and Polacheck, 1996) presents a viable alternative to the regression methods for estimation of mean school size. Probability density estimates for the bivariate (y and s) and univariate (y only) data are constructed, and under the assumption of certain detection on the trackline, mean school size may be

estimated. The method is free from assumptions about the parametric form of the detection function and can be applied to either truncated or untruncated data. Further, it is robust against changing detection conditions during a survey. Chen and Polacheck (1996) compare results of estimation by the kernel method to estimation by other (regression-based) methods. They use the same IWC/IDCR data used by Borchers (1994) in his review. The point estimates obtained seem reasonable and Chen and Polacheck (1996) state that the relatively small range of these estimates over different years and strata is indicative of the robustness of the method. Although precision is poor in comparison to a regression of $\log(s)$ against $\hat{g}(y)$, it is similar to that obtained by the previously standard IWC mean school size estimation procedure (Butterworth, 1988). Kernel estimation requires variances to be estimated using resampling techniques. This should not be seen as a limitation of the method. Indeed, the use of such techniques in regression estimation may be recommended when it is necessary to allow for estimation of the independent variable. However, the first step of the bootstrap algorithm proposed by Chen (1994) relies on strong independence assumptions and a modification of this step to use transects or days as the sampling units rather than schools might be better. The independence between school size and school location is then less critical.

Although we have only considered construction of kernels based on the bivariate detection function $g(y, s)$ here, Chen (1994) remarks that kernel estimation in higher dimensions is also possible (see Scott (1992)). However, the method is probably not amenable to inclusion of, or testing of, a large number of covariates, and in some cases is sensitive to the choice of the smoothing parameter.

3.3 Summary

The Horvitz-Thompson estimator provides the best method of estimating school size when detection on the trackline is uncertain (given that two platform data are available). Like the other methods, when only single platform data are available, it assumes that all schools on the trackline are detected, regardless of their size. It is also the only estimator which does not assume that school size is independent of encounter rate. Observable covariates are readily incorporated into the model and although variances may be calculated analytically, use of the bootstrap enables uncertainty in observed school sizes to be incorporated into variance estimates. The regression methods are easy to implement and can provide

estimates with reasonably good precision. Current research using generalized linear and generalized additive models provides the means for greater flexibility in modelling school size, and can reduce the effects of heterogeneity by including other covariates in the model. Size bias may be modelled directly using the bivariate detection function $g(y, s)$. However precision is substantially poorer compared with the regression methods. Although point estimates and variances from the kernel estimation method are comparable with those from other methods, current methodology is also restricted to kernels constructed on the bivariate detection function $g(y, s)$, with only limited potential to include other covariates.

4 Estimation by IWC Small Management Area

4.1 A Summary of Available Methods

A simple approach which facilitates abundance estimation by Small Management Area (SMA) is to calculate a prorated estimate of abundance which depends on the proportion of the area of each survey stratum that is within the SMA, assuming that the distribution of whales within blocks is uniform (Schweder *et al.*, 1993). An alternative method is taken by Butterworth *et al.* (1992) and Borchers and Chalis (1996). They define geographic strata which correspond to the areas of intersection between SMAs and survey blocks. Estimates are then obtained using standard stratified analysis methods, with the option of assuming that one or more parameters are common across strata. Difficulties with this approach are largely related to sample sizes, which are generally insufficient to permit a fully stratified analysis. Pooling across strata is subjective, but is usually necessary for reliable estimation of effective search width and may also be required for estimation of mean school size, particularly if only confirmed schools are used to estimate mean school size. (A 'confirmed school' is one whose size has been estimated with a high degree of confidence.) A third approach employs spatial models to fit a density surface across a region. Abundance estimates by SMA are obtained by integrating over the corresponding section of the surface. Using line transect data, Marques (in prep.) fits separate spatial surfaces for encounter rate and school size using generalized additive models, then (with an estimate of effective strip width which may depend on explanatory variables associated with detectability but not spatial location) combine the surfaces to obtain a density surface. By including a factor or covariate corresponding to years, spatio-temporal models can be fitted, allowing

data from more than one survey to be used. Alternative methods for modelling the spatial density surface are developed in Chapters 3, 4 and 5 of this thesis.

Chapter 3

Incorporating Spatial Variation

1 Motivation

Although it is widely recognized that whale distribution and abundance varies according to habitat and other environmental influences, such as prey distribution, conventional density estimation techniques generally rely on design-based sampling strategies. In line transect surveys, this means that estimation depends on the random placement of lines with respect to the objects of interest. Strindberg (in prep.) presents an evaluation of systematic versus random line placement in this context. Estimates from design-based surveys provide only a large-scale indication of the variation in density throughout a given survey area, with the main outputs being density and its associated variance by some geographic area or ‘stratum’ within the survey area. The main constraint on the size of the stratum is that of ensuring a sufficiently high number of sightings to facilitate analysis within that stratum.

In their report of the 1980/81 IWC Southern Hemisphere Minke Whale Assessment Cruise, Butterworth and Best (1982) implement two methods aimed at obtaining an understanding of the density distribution of minke whales (*Balaenoptera acutorostrata*) in IWC Area V with respect to the ice edge boundary. The first method, that of ‘contouring’, today seems somewhat rudimentary, and has since been superseded by more sophisticated methods. It provides hand-drawn contours indicating the estimated school density in an area, and by averaging between contours and summing throughout the surveyed area, an estimate of

the abundance of schools can be obtained. A number of shortcomings of the approach are noted by de la Mare and Cooke (1982), who indicate potential sources of positive bias. The contouring method does not provide an estimate of variance, and because of the subjective nature of obtaining the contours, cannot be easily replicated. Therefore Butterworth and Best (1982) also apply the geostatistical technique of universal kriging (Matheron (1969); see also Cressie (1991), pp. 151-183). This method assumes that a data point (in this case, estimated daily school density, which is assigned to the vessel's noon location) is a random function which can be partitioned into two components. The first component is a linear combination of known functions representing the expected value of the daily school density and characterising the large-scale trend in density; the second is a random process representing the spatial correlation and is described by a semi-variogram. The results presented in Butterworth and Best (1982) yielded abundance estimates comparable with those from other methods, but also demonstrated the need for further examination into the use of the approach for this purpose. One of the problems was that the ranges from the semi-variograms were too large to be biologically realistic. Whilst no reason for this was suggested, it seems likely that the averaging of school density over a full survey day to produce the point data necessary to apply the kriging technique serves to cloud the inherent structure of the data, and thus the spatial correlation cannot be estimated with confidence. In addition, variance estimates were negatively biased, providing only lower bounds on the true variance. The universal kriging approach for estimating whale density was not pursued further — the method being abandoned in favour of design-based estimation techniques.

Attention has recently begun to focus once more on using environmental explanatory variables to model the density distribution of a population or stock. In terms of stock management, the renewed impetus for a spatial modelling approach is based largely on its ability to provide estimates of abundance for small geographic regions that do not necessarily correspond with survey strata.

2 Introduction to Spatial Modelling in Wildlife Surveys

Before describing briefly how spatial methods have already been used to model wildlife survey data, we first outline what data are available from such surveys. Generally speaking,

spatial data comprise values of a response at known locations within a surveyed area. In wildlife surveys, the response usually indicates either the presence or absence of a species, or some countable measure of its abundance not at a point, but within a small site or 'grid square' forming part of the survey region. These data are often supplemented by additional covariate information at each location.

Cressie (1991) defines three types of spatial data which, by their nature, are suitable for different types of analysis. The first type is geostatistical data, for which kriging methods were developed, allowing modelling of the continuous spatial variability in the response. The second type is that arising from spatial point patterns, where the locations of each event are stochastic and analyses focus on examining whether a pattern exhibits complete spatial randomness, regularity or clustering. The third type — lattice data — is suitable for discrete modelling of spatial variability, and is most akin to the type of data described above for many wildlife surveys. It is nevertheless important to note that the same data set may be appropriate for more than one type of spatial analysis. In the remainder of this section, we concentrate only on one type of spatial analysis, namely regression-based methodology.

Generalized linear models (GLMs; McCullagh and Nelder (1989)) have been widely used to model species distribution using spatially-referenced covariates (for example, Osborne and Tigar (1992); Buckland and Elston (1993)), yielding informative results. More recently, greater flexibility in modelling spatial variability has been provided by generalized additive models (GAMs; Hastie and Tibshirani (1990)), and these have also been successfully applied to wildlife survey data (for example, Augustin *et al.* (1996a; 1996b); Borchers *et al.* (1997)). These studies are described briefly below.

Osborne and Tigar (1992) use logistic regression to model bird atlas data in Lesotho. The data were recorded for a number of avian species as presence or absence within one of 62 quarter-degree grid squares (each approximately 660 km²) and several habitat variables were available for inclusion in the model. Buckland and Elston (1993) also use logistic regression, modelling presence-absence of red deer (*Cervus elaphus*) in northeast Scotland. The data were again recorded by grid squares, at a resolution of 1 km². These data are also used by Augustin *et al.* (1996a) and Augustin *et al.* (1996b), who extend the logistic formulation to an autologistic model by including an autocovariate term. A second example presented by Augustin *et al.* (1996a) uses data from a survey of western mackerel

(*Scomber scombrus*) eggs in the northeast Atlantic. Unlike the datasets previously described, these data are not segregated by grid square, rather they are direct counts of the number of eggs at systematically-located sample points. By knowing the effective surface area sampled, egg densities at each sample location could be calculated. With such data, the natural model that comes to mind is one based on the Poisson distribution; Augustin *et al.* (1996a) use a GAM with an overdispersed Poisson error distribution. Borchers *et al.* (1997) take a similar approach in modelling western mackerel also in the northeast Atlantic, but employ a two-stage approach for spatial modelling of horse mackerel (*Trachurus trachurus*) in the same region, first using logistic regression to model the presence-absence of eggs from systematic sample locations, and then for those locations where eggs were present, fitting a GAM with gamma error distribution to the egg numbers.

Although the analysis methods used by these authors differ, the datasets share a common thread. The response, whether presence-absence within a grid square or a count identified as within some area, arises from a systematic survey designed to produce discrete data at specified locations. Such data contrast with data collected from a line transect survey, where the survey effort is continuous, rather than at sampled points only. The problem of estimating a (two-dimensional) spatial distribution of a species from line transect sampling may be thought of as estimating a two-dimensional surface from a one-dimensional sampler (the transect lines). This difficulty can be overcome by using a two-dimensional sampler, such as a strip transect. In the next section, an approach is developed which allows estimation of a spatial surface from strip transect data.

3 A spatial modelling formulation for strip transect data

3.1 Modelling group density

Strip transects represent a special case of line transects (Buckland *et al.*, 1993a). Transect lines of total length L are covered within a survey area A , and it is assumed that all ‘objects’ (in wildlife surveys, groups of animals) out to a perpendicular width w on either side of the lines are detected with certainty. Any detections made beyond w are excluded from the analysis.

Suppose that a strip transect survey has been conducted. The transects may be divided

into T small contiguous sampling units or ‘segments’ each of (approximately) equal length. The length of each segment should be chosen such that the geographic location does not change appreciably within a segment. (This definition is rather vague, but will suffice for now, without affecting the modelling formulation.) Let the length of the i^{th} segment be denoted by l_i and the number of groups (e.g. pods, schools, herds) detected within it by n_i , $i = 1, \dots, T$.

Now suppose that for each segment, a set of k spatial covariates is available, and let z_{ik} denote the value of the k^{th} spatial covariate in the i^{th} segment. The expected values of the n_i may be related to the spatial covariates by fitting a generalized linear or generalized additive model, the formulations of which are very similar.

As in the western mackerel example of Augustin *et al.* (1996a), the response variable is a count — here it is the number of groups detected in each segment. Using a GLM formulation with a logarithmic link function, a spatial model for strip transect sampling may be written as follows:

$$E(n_i) = \exp \left[\ln(2l_i w) + \beta_0 + \sum_k \beta_k z_{ik} \right] \quad (1)$$

where the logarithm of the area of each segment, $\ln(2l_i w)$, enters the linear predictor as an offset, and the β_j , $j = 0, \dots, k$, are parameters to be estimated. However since the error distribution has not been defined, this model is as yet incompletely specified. If the n_i are Poisson distributed, then likelihood methods can be used to fit the model. Under the Poisson assumption, the variance, $\text{var}(n_i)$, of the observations is equal to their expectation, $E(n_i)$. In general, this assumption will not be valid for transect-based wildlife surveys where sightings are often clustered, and it is often found that $\text{var}(n_i) > E(n_i)$. In this case, the data are termed ‘over-dispersed’ and quasi-likelihood methods (Wedderburn, 1974; McCullagh and Nelder, 1989) are appropriate. Rather than requiring an explicit form for the distribution of the observations, these methods require the relation between the variance of the observations and their mean to be specified.

Additional flexibility in the form of the functional dependence between the observations and the spatial covariates may be provided by the following GAM formulation:

$$E(n_i) = \exp \left[\ln(2l_i w) + \theta_0 + \sum_k f_k(z_{ik}) \right]. \quad (2)$$

Using a logarithmic link function as before, and with either Poisson-distributed observations or a specified variance-mean relationship, equation (2) is a direct extension of equation (1). Increased flexibility is provided by the (unknown) functions f_k , which are smooth functions of the spatial covariates and, analogous to the parameter vector β , need to be estimated. Thus equation (1) is a special case of equation (2), in which the f_k are linear functions.

There are two further fundamental aspects to fitting a GAM that merit brief comment here. The first involves the choice of the type of smoother to use. Possibilities include smoothing and regression splines, locally weighted polynomials, kernel and near-neighbour smoothers. However, few systematic comparisons between smoothers exist (Hastie and Tibshirani, 1990), and furthermore, there is no objective method to choose between them. Hastie and Tibshirani (1990) refer to the results in Silverman (1984) and Müller (1987) which suggest that for comparable amounts of smoothing control, there are not likely to be large differences between locally weighted running line, cubic smoothing splines and kernel smoothers. In this thesis, the objective will not be to compare smoothers — rather we present new methodology that will allow spatial models to be fitted from line transect data. The methodology may be applied using any type of smoother but without loss of generality, the examples will only use smoothing splines. This has the additional advantage of computational ease, since the smoothing spline is one of the functional forms available for use within a GAM in S-PLUS. The second component of GAM fitting is the choice as to how smooth the functions f_k should be. Smoothness is controlled by specifying the degrees of freedom (df) associated with each function, with 1 df corresponding to linearity. As the degrees of freedom are increased, the functions become progressively more flexible until at the other extreme, a fit which simply interpolates the data will be obtained if the degrees of freedom are chosen to be equal to one less than the number of data points. Usually of course, the degrees of freedom will be selected to be in between these two extremes, but more research is required to develop automated methods with a sound theoretical foundation to perform this selection. In practice, GAMs are frequently fitted using one or more of the objective measures currently available, such as Akaike's Information Criterion (Akaike, 1973) or generalized cross-validation (GCV; Craven and

Wahba (1979)), often supplemented by graphical methods. A recent advance towards a less heuristic approach to smoothing parameter selection has been made by Wood (2000), who proposes an efficient multiple-penalty approach applying GCV to this problem.

Returning now to the formulation of a spatial model for strip transect data, suppose that a GLM has been fitted to the n_i using k spatial covariates z_k . The fitted model provides estimates of the expected number of sightings within each segment of the strip. Estimated group density within each segment, \hat{D}_i , may be calculated by dividing by the area of the segment, thus $\hat{D}_i = n_i/2l_iw$ (so that $E(n_i)$ is estimated by n_i). To predict the density surface throughout the survey region, the region is divided into a grid, and with known values of the spatial covariates at the centre of each small grid square, the estimates of the parameters β_j , $j = 0, \dots, k$, are used to predict the density at the centre of each grid square. An estimated density surface is obtained by assuming that the density within each grid square is constant, and equal to that at the centre of each square. Clearly, this approximation is most reasonable when the survey region is divided into a grid of high resolution. Abundance throughout the survey region may be estimated by integrating numerically under the estimated density surface.

The general procedure for predicting from a GAM is similar to that described above for a GLM, although the non-parametric nature of the fitting (using iteratively reweighted least squares) means that it is considerably more computationally intensive. For further details, see Hastie and Tibshirani (1990).

3.2 Modelling animal density

Extending the models presented in equations (1) and (2) to estimate a density surface for individual animals is in principle very straightforward from strip transect data. A group of animals is recorded as being within the strip if the *centre* of the group is within w either side of the transect line. Assuming that the size of all groups within the strip is recorded without error, then the number of animals in the i^{th} segment, m_i , is

$$m_i = \sum_{j=1}^{n_i} s_{ij} \tag{3}$$

where s_{ij} is the size of the j^{th} detected group in segment i , and n_i is (as defined in the previous section) the number of groups in that segment. The m_i replace the n_i in the model so that for example, equation (1) becomes

$$E(m_i) = \exp \left[\ln(2l_i w) + \beta_0 + \sum_k \beta_k z_{ik} \right]. \quad (4)$$

If the size of the groups encountered is quite small, so that the number of individuals in a segment is not much greater than the number of groups, then an overdispersed Poisson error distribution, fitted by quasi-likelihood, might still be appropriate. Possible alternative error distributions, which may be more suitable when larger groups are encountered, include the negative binomial and the gamma distributions. The negative binomial distribution assumes that $\text{var}(n_i) = E(n_i) + [E(n_i)]^2/k$. The unknown parameter k (sometimes called the aggregation parameter) must be estimated from the data. Although the observations are discrete, the gamma distribution may provide a continuous approximation to the underlying error distribution. The variance function is given by $\text{var}(n_i) = \phi[E(n_i)]^2$, where ϕ is the dispersion parameter.

Finally we note that whilst strip transect surveys provide a good method for estimating the abundance of certain wildlife species (for example, feeding or resting seabirds), line transect surveys are generally preferable for species that are more sparsely distributed because the perpendicular truncation width can be larger for line transects than strip transects. Therefore line transects are much more commonly used than strip transects for cetacean surveys. In the next section, we build on the models presented above, and show how they may be extended for use with line transect data.

4 A spatial modelling formulation for line transect data

4.1 Modelling group density

The previous section showed how, within a GLM or GAM framework, spatial models could be used to estimate a density surface for a population from strip transect data, and by integrating under this surface, an estimate of abundance could be obtained. Although the modelling principles from the previous section extend readily to the line transect case, the

additional complexity induced by the need to model the detectability of the animals in question is considerable.

Suppose that a line transect survey has been conducted, and as before, the transects have been divided into T small segments such that the sighting conditions and geographic location do not change appreciably within a segment. The length of segment i , l_i , is known, and the width is given by $2w_i$, where w_i is the perpendicular sighting distance beyond which detections are excluded from the analysis (the ‘truncation’ distance). For the remainder of this section, we assume for simplicity that $w_i = w \ \forall i, \ i = 1, \dots, T$.

The sightings data comprise the number of detected groups in each segment (n_i) and the perpendicular distance of each group from the transect line, and may comprise other data specifically related to each sighted group, such as size or behaviour. The effort data record comprises the spatial information, including the geographic location of the survey effort (logged at regular intervals) and the values of the spatial covariates to be considered for inclusion in the model, and environmental factors which potentially affect the detectability of the groups (e.g. sea state, swell height).

Conventional line transect estimation requires estimation of an effective strip half-width, μ , via a detection function, g , which is a function of the perpendicular sighting distances. Recall that μ defines the width of a strip either side of the transect line within which the same number of detections is missed as is seen from μ to w . Loosely, one may envisage a line transect as a strip transect with estimated width $2\hat{\mu}$. Then, analogous to equation (1), a spatial (generalized linear) model for line transect data may be written:

$$E(n_i) = \exp \left[\ln(2l_i\hat{\mu}) + \beta_0 + \sum_k \beta_k z_{ik} \right]. \quad (5)$$

Neglecting for the moment that a component of the offset is an estimated quantity, we note that μ may be defined in terms of the truncation distance, w , and the probability of detection, p , of a group (unconditional on its position) within w (assumed to be finite) of the line: $\mu = wp$ (Buckland *et al.*, 1993a). The offset term in (5) above thus becomes $\ln(2l_iw\hat{p})$.

This has given us a simple formulation of a spatial model, given an estimate of p or equivalently μ . In its present form, it implicitly assumes that the effective strip width

is constant along the whole transect line. We now present the formulation from a different perspective, using an estimator based on the Horvitz-Thompson estimator (Horvitz and Thompson, 1952; Thompson, 1992) and show how this can lead to a model which incorporates variation in detection probabilities.

Given n_i detected groups each with detection probability p in the i^{th} segment (of length l_i and width $2w$), the true number of groups in the segment, N_i , may be estimated using as

$$\hat{N}_i = \frac{n_i}{\hat{p}}. \quad (6)$$

For example, if two groups were seen in the i^{th} segment, and they each had a 40% chance of being detected, then the estimated number of groups really present in the segment would be five. In practice, the detectability conditions might well vary in different regions of the survey area, causing the probability of detection of a group to differ spatially. We can represent this variation by introducing a segment-specific detection probability, p_i , which represents the probability of detection of a group in the i^{th} segment. Then an estimate of the number of groups in this segment is n_i/\hat{p}_i . This is analogous to the Horvitz-Thompson estimator, except that with this estimator the ‘inclusion probability’, π_i (say), is known, whereas here the corresponding probability must be estimated.

Now consider that instead of modelling the counts, n_i , directly, the response is formed by the Horvitz-Thompson-like estimates of the number of groups in each segment. Applying a logarithmic link function and using the GLM formulation as above, the model may be written as either

$$E(\hat{N}_i) = \exp \left[\ln(2l_i w) + \beta_0 + \sum_k \beta_k z_{ik} \right] \quad (7)$$

or

$$E(n_i) = \exp \left[\ln(2l_i w \hat{p}_i) + \beta_0 + \sum_k \beta_k z_{ik} \right]. \quad (8)$$

Although the two forms are algebraically equivalent, the error distributions will be differ-

ent. For equation (8), a Poisson distribution might be appropriate. If not, and overdispersion is present, quasi-likelihood methods can be used as described in section 3.1. The response \hat{N}_i in equation (7), although derived from count data, would not be expected to follow a Poisson distribution even if the counts themselves were Poisson—overdispersion having been induced by adjusting the counts according to their probability of detection. Again, either quasi-likelihood methods or an alternative error distribution would likely be more appropriate.

There was, however, a reason for presenting and demonstrating the equivalence of the two forms (7) and (8) which will now become clear. If we are to incorporate the additional individual variation in detection probability of each group (caused, for example, by markedly different behaviour or differences in group size) within each segment, then it is necessary to use the first of the two forms (i.e. equation (7)) and model the Horvitz-Thompson-like estimates directly. In this case, N_i is estimated by

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{1}{\hat{p}_{ij}}, \quad (9)$$

where \hat{p}_{ij} is the estimated probability of detection of the j^{th} detected group in segment i .

4.1.1 The Horvitz-Thompson-like estimates: estimating p , p_i or p_{ij}

If it is reasonable to assume a constant detection probability, p , for all groups, then estimation is straightforward and conventional line transect analyses, which yield an estimate of p , may be used. Software such as Distance (Thomas *et al.*, 1998) facilitates such an analysis. If detection probability varies spatially, or varies spatially and with each individual group, then p_i or p_{ij} must be estimated. In the case of two-platform survey data, this can be achieved by using logistic regression as in Borchers *et al.* (1998). In the case of single platform data, the covariate information may be incorporated by adopting a full likelihood approach (Marques, in prep.; Cooke and Leaper, 1998; Ramsey *et al.*, 1987), or by using a ‘covariate adjustment’ method (Beavers and Ramsey, 1998), in which a log-linear regression is carried out on the observed perpendicular distances to modify the width of the effective search area dependent upon the sighting conditions.

4.2 Modelling animal density

In section 3.2, we presented a model for estimating animal density from strip transect data. This involved modelling the number of animals directly (rather than the number of groups of animals). The model was based on the assumptions that all groups whose centres are within the strip are detected with certainty, and further, that the number of animals in each group is recorded without error. With line transect surveys it is not assumed that all groups are detected and instead the detectability of the groups is estimated, customarily by modelling the decrease in detection probability with increasing distance from the transect line. In addition, estimation of mean group size should take into account any variation in detection probability caused by large groups being relatively more conspicuous than small groups (the ‘size bias’ effect). A number of methods already exist to estimate mean group size reliably (see Chapter 2, section 3). Therefore one way to obtain estimates of animal density is to use one of these methods to estimate mean group size in the surveyed region (or, if appropriate, estimate mean group sizes separately for strata within the surveyed region), then simply multiply the estimate(s) by the corresponding region(s) of the group density surface obtained by one of the models described in section 4.1. An alternative approach is to adopt a strategy similar to that described in section 3.2 and model the estimated number of animals in each segment directly. This method still relies on the second of the two assumptions underlying the strip transect model — namely that the group sizes are recorded without error. In practice with many cetacean species, estimation of group size is very difficult, and on many surveys, groups are recorded as either ‘confirmed’ or ‘unconfirmed’, where a confirmed group is one whose size has been determined with a high degree of confidence. Thus, if the confirmed group size is available for every detected group, the estimated number of animals in the i^{th} segment, \hat{M}_i , may be modelled. For example, if the probability of detection in each segment is estimated separately for each school, then M_i is estimated by

$$\hat{M}_i = \sum_{j=1}^{n_i} \frac{s_{ij}}{\hat{p}_{ij}}, \quad (10)$$

where s_{ij} is the confirmed group size for the j^{th} school in the i^{th} segment. Since one of the covariates used to model \hat{p}_{ij} is likely to be group size, the method accommodates size bias in detection of groups. A spatial (generalized linear) model for animal density may

thus take the following form:

$$E(\hat{M}_i) = \exp \left[\ln(2l_i w) + \beta_0 + \sum_k \beta_k z_{ik} \right]. \quad (11)$$

Since any overdispersion in the raw observations (the number of schools in each segment) relative to a Poisson distribution will have been greatly exacerbated by forming the response in this way, a negative binomial or gamma error distribution may be appropriate, although the model could still be fitted by quasi-likelihood assuming an overdispersed Poisson error distribution.

In some circumstances, it may be preferable to model the group size separately, obtaining separate spatial surfaces for the group size and the group density. In principle these could then be multiplied together to obtain a surface for animal density. Statistically, this method presents no difficulties; a GLM or a GAM with, for example, a gamma error distribution and logarithmic link function to ensure positive fitted values, may be fitted to the observed school sizes. However, it is not clear how to account for size bias with such an approach and so in practice this may often be an inadequate solution.

5 Variance estimation

5.1 The problem of dependent observations

In section 3, we introduced a spatial modelling formulation for strip transect data based on dividing the transects covered into smaller sampling units ('segments'), then in section 4, it was shown how this formulation could be extended for line transect data. The spatial model itself was based on either a GLM or a GAM, and the response variate was either the number of groups, or the number of animals, in a segment. Both GLMs and GAMs treat the observations of the response as being independent random variables. However, this independence assumption is clearly invalid here, since the counts from neighbouring segments are correlated. Three implications of the dependence of the segment counts are:

1. Goodness-of-fit statistics based on the χ^2 distribution will have artificially-inflated degrees of freedom.

2. Theoretical variance estimators will be biased.
3. Model selection criteria based on the likelihood function such as Akaike's Information Criterion (AIC; (Akaike, 1973)) will be invalidated.

When comparing nested models, the first of these problems has little practical consequence since the goodness-of-fit is then based on the *difference* in deviances rather than the absolute deviances. The second problem is overcome by using resampling methods to estimate variance. In the next section, three such approaches are described: the non-parametric bootstrap, the jackknife and the parametric bootstrap (Davison and Hinkley, 1997; Efron and Tibshirani, 1993). The extent of the difficulties encountered by the third problem, the invalidation of likelihood-based model selection methods, depends to some degree on the modelling objectives. For example, AIC will tend to select a more complex model when the independence assumption is invalid. Such overfitting would not be expected to affect substantially the estimation of density, but inferences about the relationships between the response and the explanatory variables would be unreliable.

5.2 The non-parametric bootstrap

The non-parametric bootstrap is usually applied to a set of n observations by selecting samples of size n from the observations at random and with replacement. It is based on the assumption that the observations are independently and identically distributed (iid), which in line transect surveys may not be a valid assumption, and in the spatial modelling context where the observations are counts per segment, almost certainly is not. In both cases, more reliable estimation of variance should be achieved by using the transect lines, or effort corresponding to individual days, as sampling units (since it is more reasonable to expect these to be iid).

However if the non-parametric bootstrap is applied with the transect lines (say) as the sampling units, then because the resamples are drawn with replacement, each bootstrap resample is likely to generate poor spatial coverage relative to the original survey effort, particularly when the number of transect lines is small. Therefore although the non-parametric bootstrap *may* be suitable in some circumstances, perhaps where there is good survey coverage and a large number of transects — as, for example, in the 1994 SCANS

survey (Hammond *et al.*, 1995, see Figure 1) — it is not considered in further detail here.

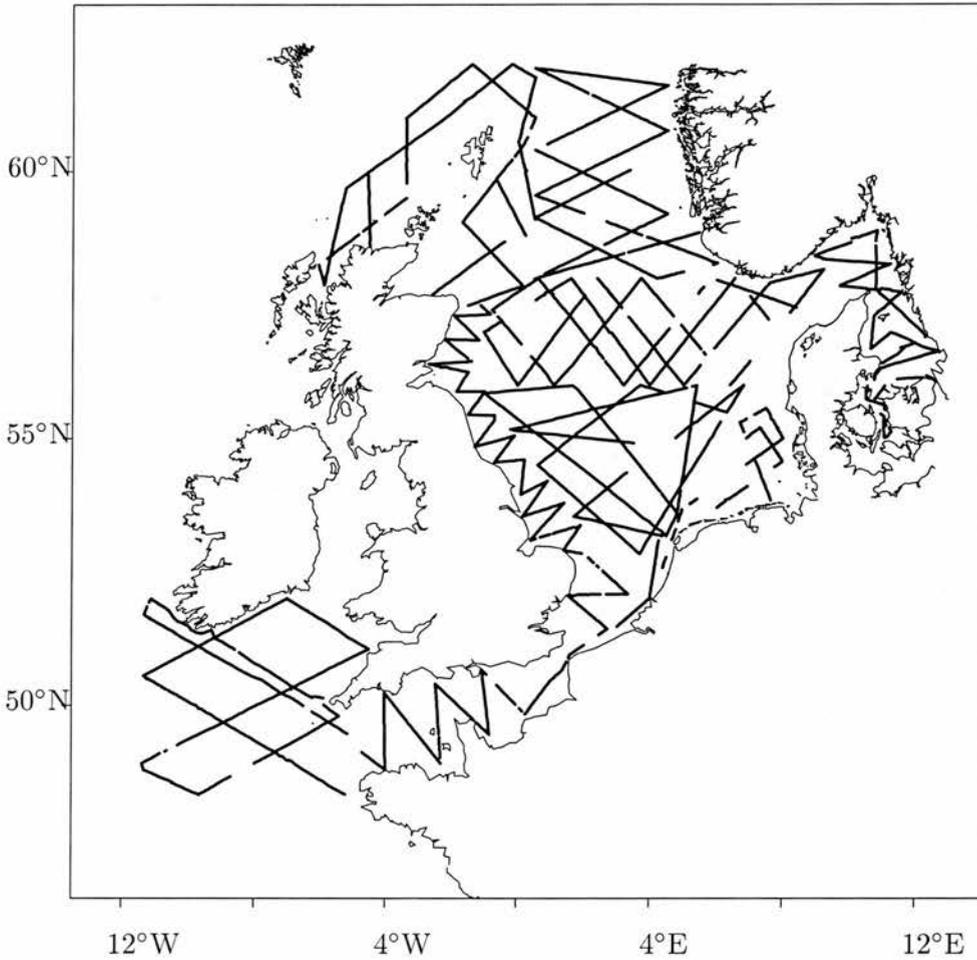


Figure 1: Realized shipboard survey effort from the 1994 SCANS survey of the North Sea and adjacent waters. The large number of transects and good coverage mean that the non-parametric bootstrap could be a suitable method for estimating variance.

5.3 The jackknife

The jackknife is usually applied to a set of n observations by systematically removing one observation at a time to create n samples of size $n - 1$.

Although the jackknife could be implemented in this way here, by removing one detection (not one segment) at a time, it is more useful to define the sampling units as the transect lines (or days of survey effort) as for the non-parametric bootstrap above (Buckland *et al.*, 1993a). The jackknife samples are then obtained by systematically deleting one transect (or day) at a time from the observed data. Suppose that the sampling units are taken

to be the lines and there are τ of these in the original data, then each of the τ jackknife samples are analysed as if they were the real data.

For the line transect models described in section 4, this first involves re-estimation of the probability of detection of each school, whether this is the simple case with a constant probability, p , for all schools, a constant probability, p_i , for each school within a segment, or a different probability, p_{ij} , for each school within a segment. Taking the last case for example, we denote the estimates of the probability of detection of the j^{th} school in the i^{th} segment as $\hat{p}_{ij}^{(-t)}$, $t = 1, \dots, \tau$, where the $(-t)$ superscript indicates that the estimate is calculated from the t^{th} jackknife sample (ie. with the t^{th} transect omitted). A spatial model is then fitted to the Horvitz-Thompson-like estimates of the number of schools in each segment for each of the τ jackknife samples. (Model selection uncertainty could be incorporated at this stage by refitting the GLM or GAM for each pseudosample. In some circumstances, it may be desirable to exclude this source of uncertainty — perhaps for comparison purposes with conventional analytic estimates for example — in which case one should condition on the original selected model for the real data and fit that to each pseudosample.)

Estimates of abundance from the jackknife samples, $\hat{N}^{(-t)}$, $t = 1, \dots, \tau$, may be calculated by integrating under the density surface predicted from the respective model fits.

The jackknife estimator of the variance of the abundance estimate (from the original data), \hat{N} , is calculated by first constructing τ so-called *pseudovalues*, $\tilde{N}^{(-t)}$, $t = 1, \dots, \tau$, which are defined as follows:

$$\tilde{N}^{(-t)} = \tau\hat{N} - (\tau - 1)\hat{N}^{(-t)}, \quad t = 1, \dots, \tau. \quad (12)$$

The idea behind equation (12) is that by inflating the variability between the pseudovalues, they are supposed to act as if they were τ independent values (Efron and Tibshirani, 1993). The variance of \hat{N} is then:

$$\text{var}(\hat{N}) = \frac{\sum_{t=1}^{\tau} (\tilde{N}^{(-t)} - \tilde{N}^{(\cdot)})^2}{\tau(\tau - 1)}, \quad (13)$$

where $\tilde{N}^{(\cdot)}$ is the mean of the abundance from the pseudovalues. Equation (13) is thus

the sample variance of the pseudovalues divided by τ .

5.4 The parametric bootstrap

For the parametric bootstrap, instead of resampling from the observations, resamples are taken from the model fitted to the original data.

Thus, in this spatial modelling context, the first step is to fit the GLM or GAM model to the original data. The density of groups at every point along the trackline may then be estimated from the fitted model (in practice, density is predicted along the trackline for a set of closely-spaced points). For each bootstrap resample, a single deviate, \tilde{n} , from the Poisson distribution with rate equal to the expected total number of detections in the original data, $E(n)$ (approximated by the total number of detections in the original data, $n = \sum_{i=1}^T n_i$), is generated. A rejection-sampling method is then used to generate \tilde{n} groups in the resample, as follows:

1. Generate a deviate from a uniform distribution, $\text{Unif}(0, \hat{D}_{max})$, where \hat{D}_{max} is the maximum predicted density at a location along the trackline.
2. Generate a deviate from a uniform distribution, $\text{Unif}(0, L)$, where $L = \sum_{i=1}^T l_i$ is the total length of the trackline.
3. The two uniform deviates obtained define a point on a two-dimensional co-ordinate system (\hat{D}, l) . If this point is 'below' the fitted density curve, then it is accepted; otherwise, it is rejected (Figure 2).
4. Continue until \tilde{n} points have been accepted.
5. Using the accepted $\text{Unif}(0, L)$ deviates, calculate the number of detected groups in each segment of the trackline.

As described above for the jackknife estimator, a GLM or GAM is then fitted to each resample, either conditioning on the original model fit, or refitting the model each time to incorporate model selection uncertainty. An estimate of abundance from each of the resamples may be calculated as described before. The variance of \hat{N} comprises two components: the first is due to estimation of the p_{ij} , and the second is due to the spatial modelling.

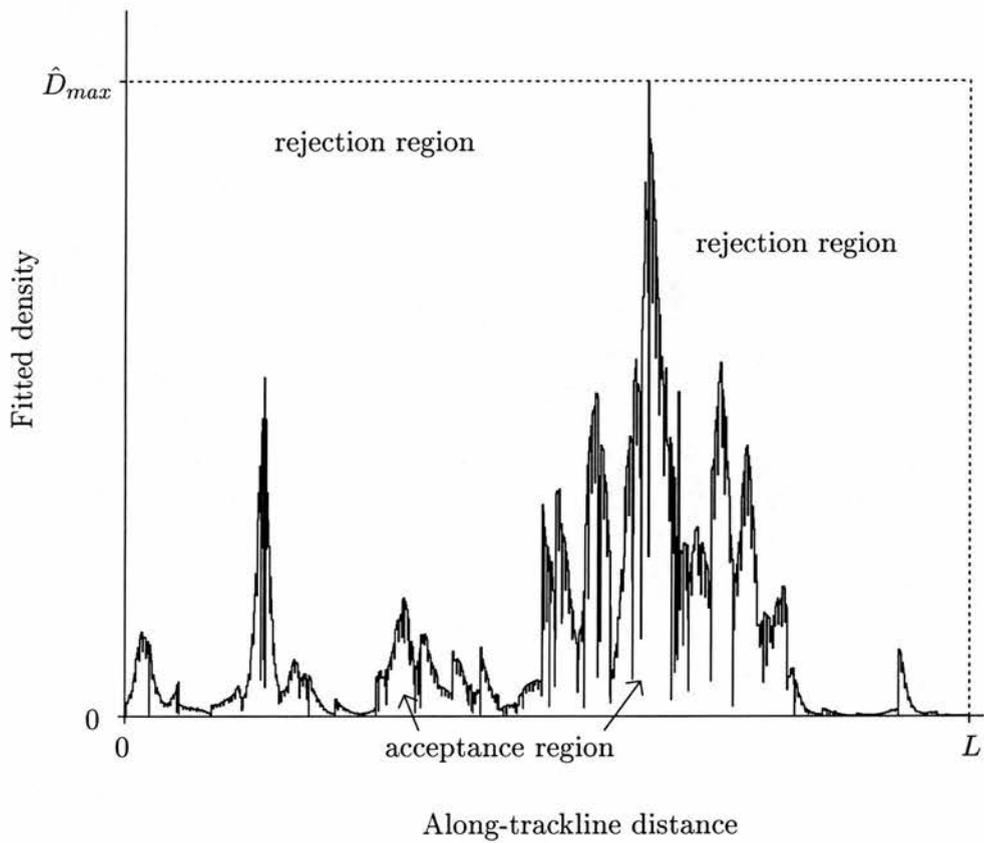


Figure 2: Density plotted against the along-trackline distance for estimation of parametric bootstrap resamples. Uniform deviates generated from $\text{Unif}(0, \hat{D}_{max})$ and $\text{Unif}(0, L)$ define a point in the area enclosed by the dashed lines. A point is accepted if it falls below the density curve, and is rejected otherwise.

The sample variance of the abundance estimates from the parametric bootstrap resamples is an estimate of the component of variance of \hat{N} due to the spatial modelling. The delta method (Seber, 1982) may be used to combine this component of the variance with the component of the variance due to estimation of the p_{ij} , and thus obtain an estimate of the overall variance of \hat{N} .

The advantage of the parametric bootstrap is that it does not assume that the observations are iid. Unlike the jackknife, where the number of resamples is determined by the number of sampling units, the number of bootstrap resamples to generate is restricted only by computer power/time, but a minimum of 200 should probably be generated for reliable estimation of variance.

6 Applications of the methodology

6.1 Line transects: a shipboard survey of Antarctic minke whales

In this example, we apply the method described in section 4 to line transect data from the 1992-3 IWC/IDCR Antarctic minke whale survey. The survey was conducted in two survey modes: Closing mode and Independent Observer (IO) mode. Here we use the data from the IO mode transects only. These are shown in Figure 3, together with the locations of detected groups.

Following Borchers and Cameron (1995), a conventional stratified analysis was conducted in DISTANCE (version 2.2) (Laake *et al.*, 1996) to estimate the probability of detecting a minke whale pod within a truncation width w of the trackline. As in Borchers and Cameron (1995), data were truncated at 1.5 nautical miles (n.miles), and effective strip widths were calculated separately for the northern strata (WN and EN) and the southern strata (WS and ES). As shown in Table 1, there is a considerable difference between the estimates from each stratum, and furthermore, the differences are counter-intuitive. In the northern strata, the weather conditions encountered are generally poorer than in the southern strata and so it would be expected that the estimated effective strip width for the northern strata would be smaller than that for the southern strata. However, Ensor (pers. comm.) has noted that it is common for the *Shonan Maru* (the vessel which covered the WN and ES strata) to have a greater estimated effective strip width than that of

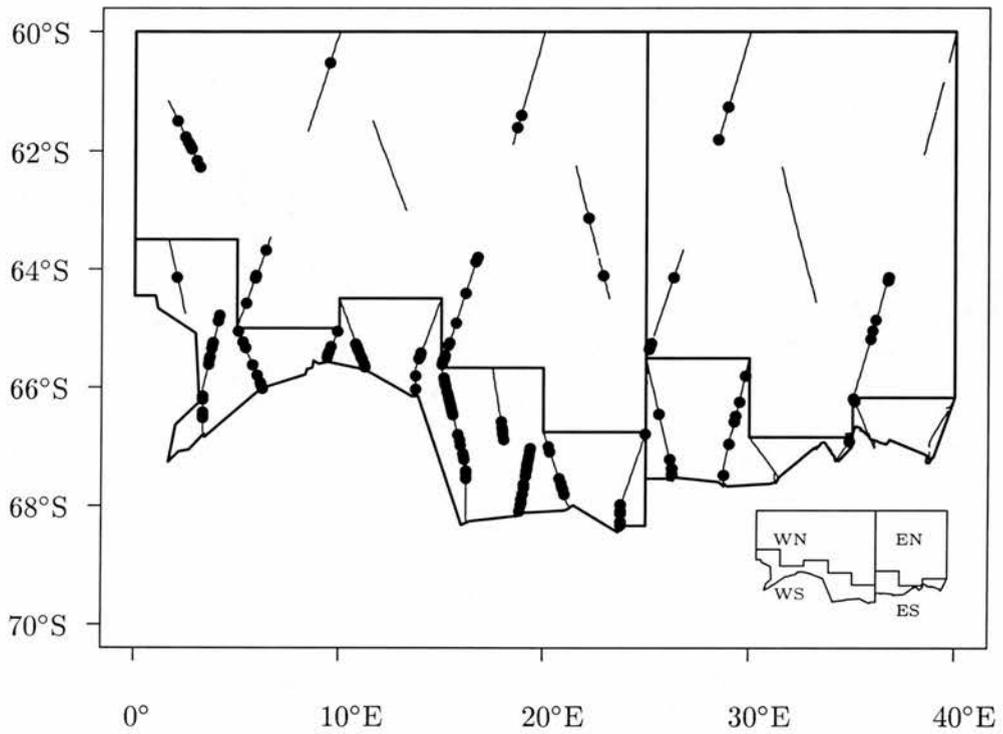


Figure 3: Realized survey effort in IO mode and minke whale school sightings during the 1992-3 IWC/IDCR Antarctic Survey. The southern survey boundary is defined by the extent of sea ice from the Antarctic continent. Subplot shows the division of the region into four strata: WN, WS, EN and ES.

the *Shonan Maru No. 2* (which covered the WS and ES strata), despite the two vessels being very alike. As yet, no satisfactory explanation for this has been found, but it seems possible that some sort of vessel effect may have produced the counter-intuitive estimates seen in Table 1. We do not pursue this issue further, the primary purpose of this section being to demonstrate an application of the spatial modelling methodology proposed in this chapter.

Pooled strata	\hat{p}	%CV	$\hat{\mu}$ (n.miles)	%CV
WN and EN	0.742	7.61	1.112	7.61
WS and ES	0.360	15.42	0.540	15.42

Table 1: Stratum estimates of detection probability, \hat{p} , of groups within the strip of half-width 1.5 n.miles, and effective strip half width, $\hat{\mu}$, with coefficients of variation (calculated using DISTANCE software).

For the spatial model, survey effort was divided into segments of 16 minutes, or approximately 3 n.miles assuming a vessel speed of 11.5 knots. No attempt was made to ensure that all segments were exactly this length — segments at the end of legs were generally shorter, but this was accounted for by including the length of each segment as a component of the offset term in the model. There were a total of 1062 segments, of which about 70% were exactly the same length. The number of detections in each segment was calculated, and the values of the response variable, \hat{N}_i , $i = 1, \dots, 1062$, were estimated as

$$\hat{N}_i = \frac{n_i}{\hat{p}_i}, \quad i = 1, \dots, 1062. \quad (14)$$

In this example, \hat{p}_i is the appropriate one of the two \hat{p} 's in Table 1. The spatial covariates available for inclusion in the model were: distance from the ice edge boundary (*icedge*); latitude (*lat*); and longitude (*lon*). Each of these was considered for inclusion in the model as a cubic smoothing spline with either 8, 4 or 2 degrees of freedom, or as a linear term. Adopting an automated stepwise procedure which uses a version of AIC that adjusts for overdispersion (Chambers and Hastie, 1993), the model selected was highly flexible with all three covariates selected as smoothed terms with eight degrees of freedom, as shown below:

$$E[\hat{N}_i] = \exp [\ln(2l_i w) + \theta_0 + s(\text{icedge}_i, 8) + s(\text{lat}_i, 8) + s(\text{lon}_i, 8)] . \quad (15)$$

The nonlinear form of the dependence of group density on the covariates is shown in Figure 4. In interpreting the plots in this figure, it is important to note that they show the additional effect of the covariate being plotted, given that the other (smoothed) covariates are included in the model. For example, the second peak in density at around 62°S in the smoothing spline of latitude does not correspond to the region of highest density of the predicted density surface (Figure 5) because, as can be seen from the smoothing spline of distance from the ice edge, densities generally decrease with increasing distance from the ice edge boundary. This was located considerably further south than (and hence at large distances from) 62°S during this survey.

The predicted density surface of minke whale schools (Figure 5) is reasonably well supported by the observed data (Figure 3), indicating that the model has provided a good description of the spatial variation in group density. The highest density area is located around 66°S, 16°E. However, a moderately high density patch is also predicted to occur at around 20°E, just north of 66°S, on the interstratum boundary between the WN and WS strata which, since there were no transects in this region, is not apparent from the sightings data. It is easy to see from Figure 4 why high densities are being predicted here. The middle plot, showing how density varies with longitude, displays a global peak at about 20°E, whilst the top plot indicates a local peak at around 66°S. These two effects together with relatively high densities predicted from the distance from ice edge smooth combine to produce this patch. (The centre of the patch is approximately 125 nautical miles from the ice edge boundary.)

The ‘truth’ is of course unknown in this example, but the scenario serves to remind us of a possible pitfall when combining one-dimensional smoothers to produce a two-dimensional surface. An alternative is to model the surface directly using a bivariate smoothing function, although the increase in complexity can lead to difficulties in interpretation and computation (Hastie and Tibshirani, 1990).

By integrating under any area of the predicted density surface of Figure 5, an estimate of abundance in that area can be obtained. For comparison with the conventional line

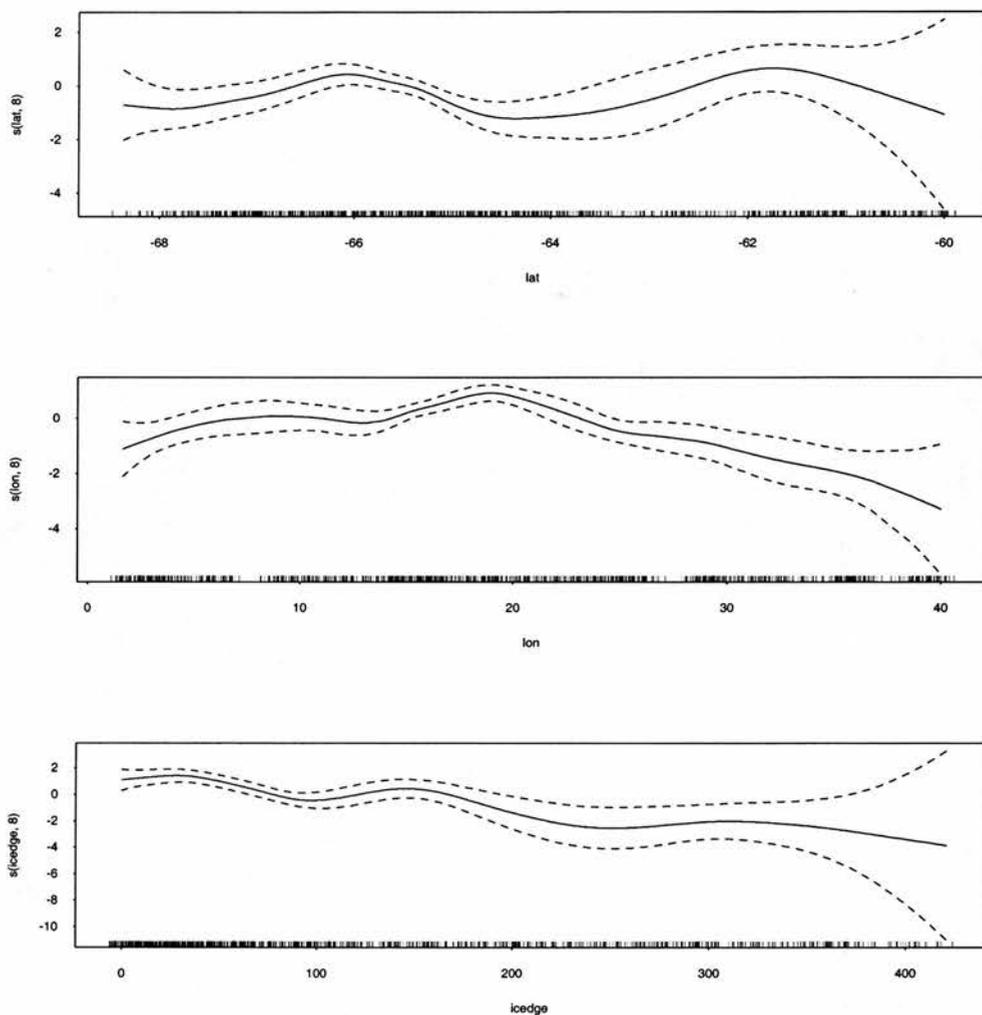


Figure 4: Shapes of the functional forms of the smoothed covariates (unbroken lines) used in the model described in equation 15. Dashed lines show standard error bands. Zero on the vertical axes corresponds to no effect of the covariate on the estimated response (group density). The locations of the observations (segments) are plotted as small ticks along the horizontal axes.

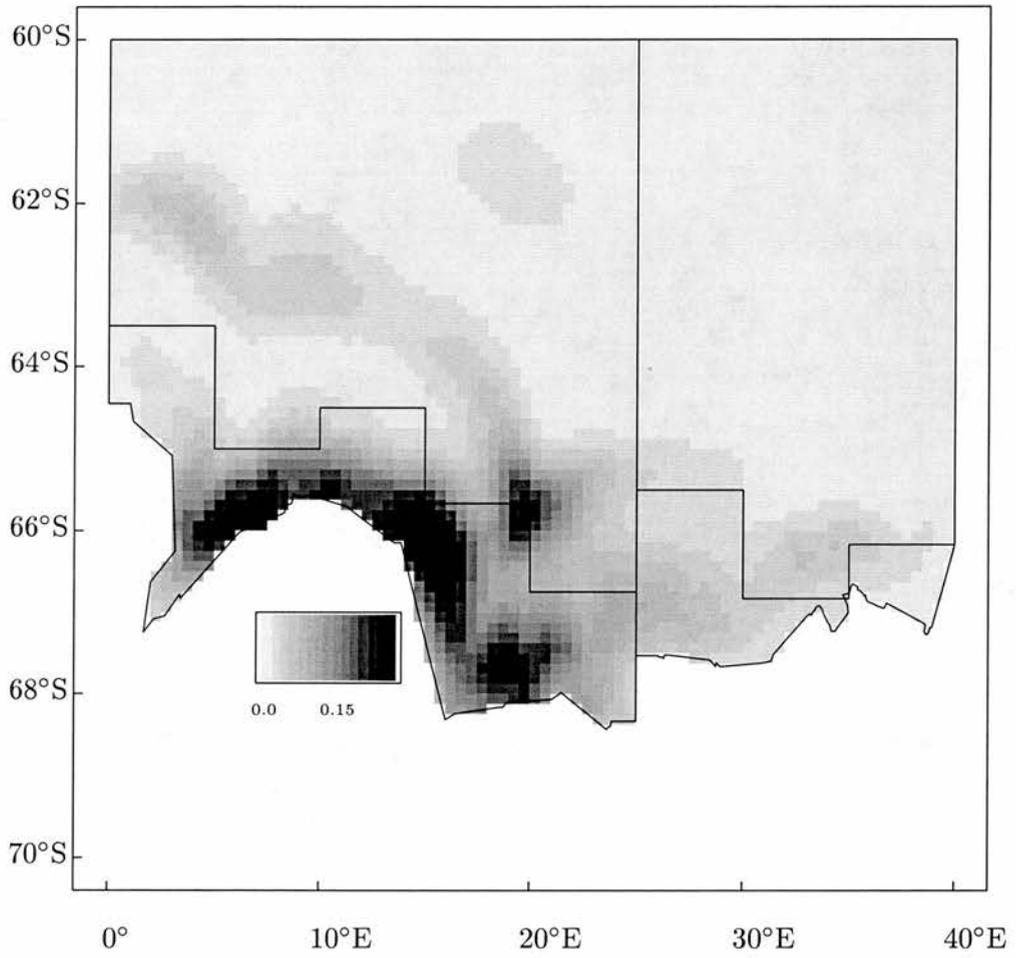


Figure 5: Density of minke whale schools in the surveyed region, predicted from a model based on cubic smoothing splines of distance from the ice edge, latitude and longitude.

Stratum	Stratified Analysis		Spatial Model		
	\hat{N}	%CV	\hat{N}	%CV _J	%CV _{PB}
WN	4,810	40.1	4,620	40.4	19.9
EN	1,460	49.5	820	44.4	28.9
WS	7,410	25.1	8,420	30.6	17.6
ES	640	44.3	880	43.3	25.1
TOTAL	14,320	23.0	14,740	25.5	16.0

Table 2: Comparison of estimates of abundance of minke whale schools (\hat{N}) from a conventional stratified analysis (Borchers and Cameron, 1995) and the spatial modelling approach. For the spatial model, the CVs were estimated using the jackknife (denoted %CV_J) and the parametric bootstrap (%CV_{PB}).

transect analysis, estimates of abundance by stratum are shown in Table 2, together with coefficients of variation (CVs). The CVs were estimated using two of the estimation techniques described in section 5, the jackknife and the parametric bootstrap, in both cases conditioning on the model fitted to the original data for each resample.

Although there are differences between the stratum estimates from the stratified analysis and the spatial model (Table 2), the total estimates of abundance from the two methods are very similar, the point estimate from the spatial model being about 3% greater than that from the stratified analysis.

In this example, the CV for the total estimate of abundance from the stratified analysis is slightly better than that estimated by the jackknife. The CVs for abundance in the WN and ES strata vary little between the analytical estimate from the stratified analysis and the jackknife, but the analytical CV for the estimate of abundance in the WS stratum is lower than that estimated by the jackknife. The WS stratum had 70% of the 224 sightings. With such a sample size, the analytical CV would be expected to perform well. In contrast, only 12 sightings were made in the EN stratum, and estimated precision is correspondingly low, although the jackknife provided improved precision over the analytical estimate. This result highlights a merit of the spatial modelling approach: improvements in precision can be expected for strata with few sightings. The adoption of a model-based rather than a sampling-based approach means that data from outwith a stratum can still provide

information relevant to the estimation of density within it, provided that an appropriate model has been used.

In contrast to the jackknife estimates, the parametric bootstrap CVs provide an improvement in precision to both the stratum estimates and the total abundance estimate. While this result does appear to be promising, the estimates may well be negatively biased — this bias having been induced by failing to incorporate fully the along-transect correlation between observations in constructing the bootstrap resamples. The estimated dispersion parameter from the original data was 2.77, much higher than the mean of the dispersion parameter estimates from the bootstrap samples, 1.68, providing evidence that the resamples do not fully reflect the clustering present in the original data. A Monte Carlo test, carried out by assessing how extreme the value of 2.77 was compared with the estimates from the resamples, indicated that the two values were significantly different (p -value=0.02). Clearly, further research is required, either to estimate the extent of the bias, or to develop methodology which is better able to incorporate the correlation structure.

6.2 Cue counts: an aerial survey of minke whales off West Greenland

In this chapter, we have focussed on applying spatial modelling methodology to line transect and strip transect data. This example shows how the approach can be applied to cue-count data (Hiby and Hammond, 1989). The data are from an aerial survey conducted off West Greenland in the summer of 1993. Larsen (1995) reports estimates of abundance of minke and fin whales resulting from this survey.

The surveyed tracklines and locations of minke whale cues (those sighted during on-effort periods only) are shown in Figure 6. The survey region comprised six survey blocks and because previous surveys had indicated that higher densities of minke whales occurred nearer the coast, all but the two most northerly blocks (1A and 1B) were subdivided into coastal and offshore strata. Effort was prioritized such that more time was spent in the coastal strata, and also higher priority was given to the two southernmost blocks (4 and 5) (Larsen, 1995). Out of a total 58 on-effort cues, 46 (79%) occurred in the coastal strata of these two blocks, and most of these occurred within 12 n.miles of the coastline. The extreme variation in the density of minke whales throughout the survey area is difficult to model using standard cue-counting methodology because small sample

sizes in the majority of survey blocks preclude a fully stratified analysis.

Because the data had undergone further validation and modification since the original analysis by Larsen (1995), a re-analysis of the data using standard cue-counting methodology was presented in Hedley *et al.* (1997b), the results of which we use for comparison with a spatial modelling approach.

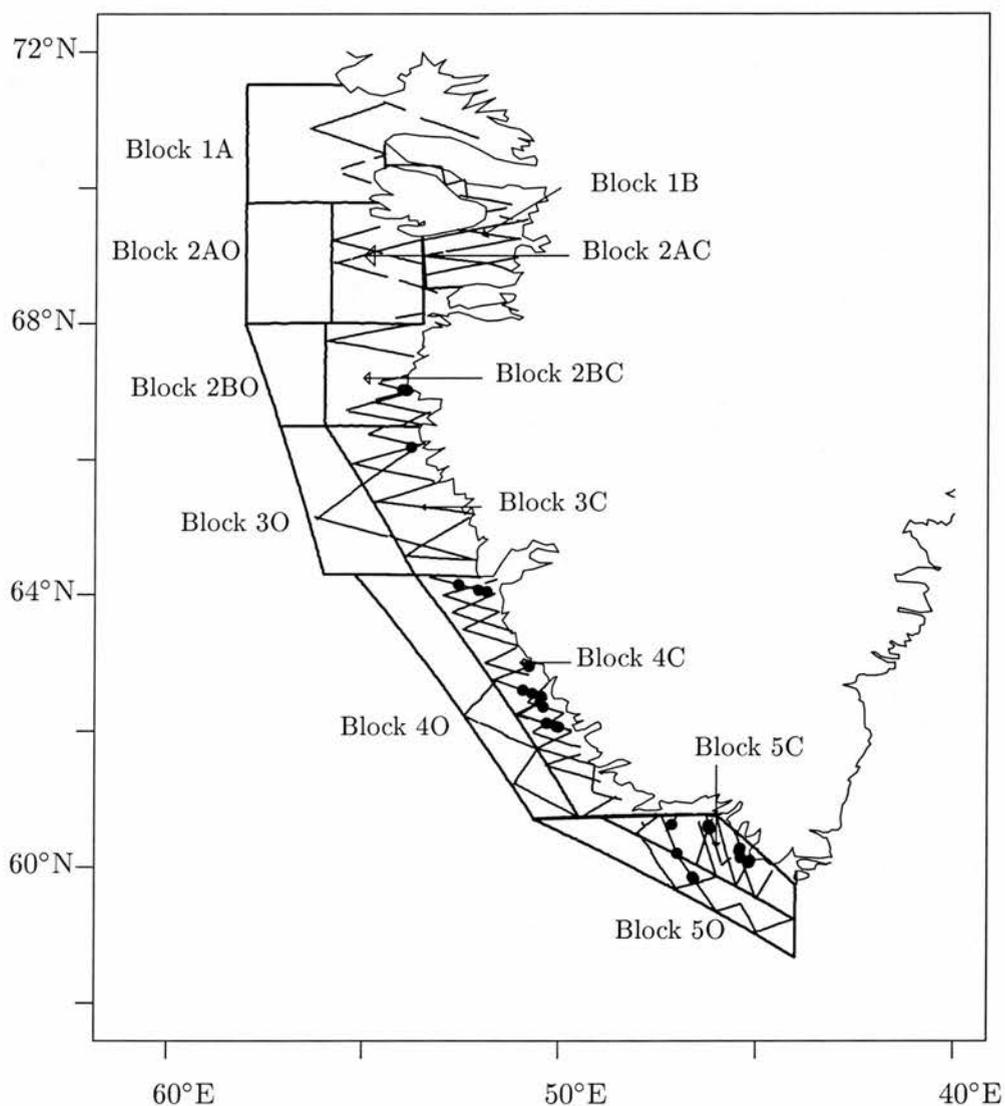


Figure 6: Maps of survey area showing block boundaries, realized survey effort and positions of all sighted cues.

Following previous analyses of these data (Larsen, 1995; Hedley *et al.*, 1997b), all effort in sea states greater than Beaufort 3 was discarded from the spatial analysis. Thus transects covered on effort in sea states 0, 1 or 2 were split into small segments of 0.5 minutes, or approximately 0.75 n.miles assuming an aircraft speed of 90 knots. This distance

is approximately equal to the effective detection diameter calculated from the stratified analysis of Hedley *et al.* (1997b). Thus the total number of segments was 5557 of which about 70% represented exactly equal amounts of time on effort. The estimated probability of detection of a particular cue, \hat{p} , within a semi-circle of radius 1.3 km (the distance at which radial distances were truncated in Hedley *et al.* (1997b)) was 0.175 (%CV 26.06).

The covariate adjustment method proposed by Beavers and Ramsey (1998) which accommodates variability in detection probability caused by differing sighting conditions was successfully implemented by Fancy (1997) for variable circular-plot counts to estimate bird densities. We investigated the use of this approach for these data, looking at the effects of ice cover, cloud cover, sea state and observer on detectability. Although it is likely that one or more of these variables affected cue detectability, the small sample sizes resulted in no significant effects being detected. Therefore the probability of detection of the j^{th} cue in the i^{th} segment was estimated by $\hat{p} \forall i, j$ and the following form of generalized additive model was fitted to the number of cues, c_i , detected in each segment:

$$E(c_i) = \exp \left[\ln \left(\frac{\pi w^2 t_i \hat{p}}{2} \right) + \theta_0 + \sum_k f_k(z_{ik}) \right], \quad (16)$$

where the offset is the logarithm of the product of the area of the (semi-circular) segment, $\pi w^2/2$, the probability of detection of any one cue in that segment, \hat{p} , and the time spent searching that segment t_i . An overdispersed Poisson error distribution was assumed, with variance proportional to the mean.

Three spatial covariates, z_k , were available for inclusion in the model: latitude (*lat*), longitude (*lon*) and the closest distance from the coastline (*coast*). In this example, model selection was carried out with a specific objective in mind, namely to obtain a predicted surface which represented the underlying distribution of minke whale cues in the survey region, rather than one which was able to predict ‘hot-spots’ of density which may have occurred at the time of survey. We therefore chose an initial nominal value of 4 degrees of freedom for the smooth of each covariate and the first stage of the model selection was to decide how many of the spatial covariates should be included in the model. As in section 6.1, this was done using a version of the AIC that adjusts for overdispersion (Chambers and Hastie, 1993), and all three covariates were selected. The second component of the model selection procedure was to evaluate whether each covariate should be included as

a smoothing spline with 4 or 2 degrees of freedom, or as a linear term. As mentioned in section 5.1, reliance on the AIC for model selection may lead to overfitting and so graphical methods were used to help select the number of smoothing parameters for each term. Although this inevitably introduces some subjectivity into the model selection, it is an approach that is widely used in practical applications of generalized additive model fitting, particularly when several smoothing parameters need to be selected simultaneously (Hastie and Tibshirani, 1990). The form of the final model was:

$$E(c_i) = \exp \left[\ln \left(\frac{\pi w^2 t_i \hat{p}}{2} \right) + \theta_0 + s(\text{coast}, 4) + s(\text{lat}, 4) + \text{lon} \right]. \quad (17)$$

The nonlinear form of the dependence of cue density on the smoothed covariates, *coast* and *lat*, shows density generally decreasing with increasing distance from the coastline and with increasing latitude (Figure 7). As with conventional cue-counting methodology, to estimate the density of minke whale groups from the predicted cue density, additional information on the surfacing behaviour of the species is required. Following Hedley *et al.* (1997b) and as originally reported in Hiby *et al.* (1989), a rate of 53 surfacings per hour was assumed. Note that no variance is attached to this estimate. Hence, the predicted density surface of minke whale groups was estimated, and its logarithm (which in this case, better illustrates the spatial variation), is shown in Figure 8, for those survey blocks in which cues were detected. A region of high density is predicted to occur along the coastline, with highest densities occurring in the southernmost coastal block, 5C.

By integrating under the appropriate areas of the predicted density surface, estimates of abundance by stratum were obtained. Coefficients of variation were calculated using the jackknife and the parametric bootstrap, and as in section 6.1, because the purpose of the example is a comparison with conventional methodology, there was no attempt to incorporate model selection uncertainty. The results are shown in Table 3.

Examination of the point estimates reveals a potential merit of the spatial approach, in that it is better able to model trends in density than the stratified analysis. With the spatial model, minke whale density is highest in the most southerly block (5C), and decreases with increasing latitude. This smooth trend is in contrast with the stratified analysis which fits a step function with uniform density within strata. Block 3C for example, in which only

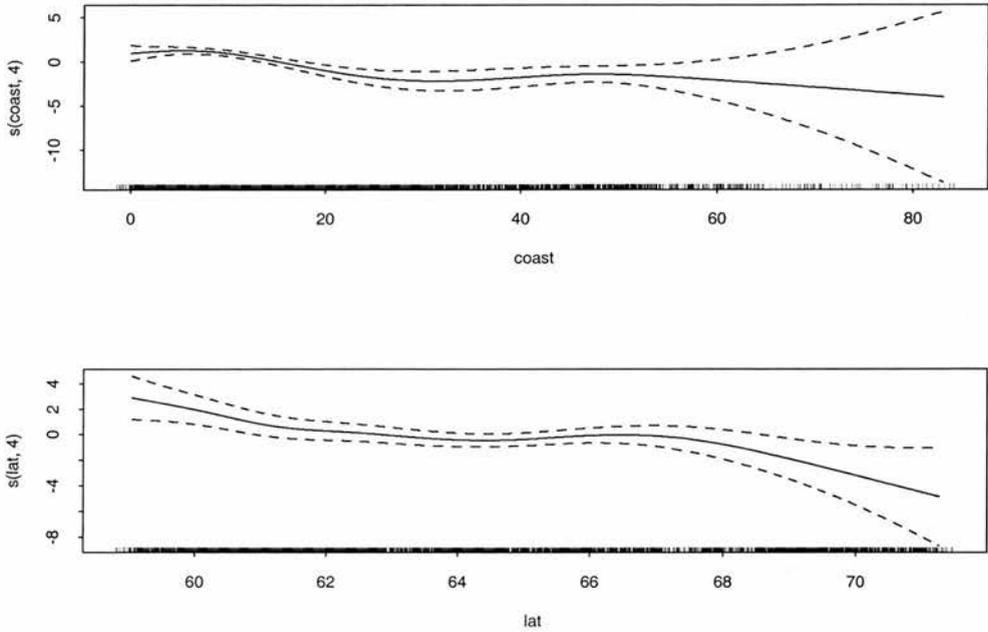


Figure 7: Shapes of the functional forms of the smoothed covariates (unbroken lines) used in the model described in equation 17. Dashed lines show standard error bands. Zero on the vertical axes corresponds to no effect of the covariate on the estimated response (cue density). The locations of the observations (segments) are plotted as small ticks along the horizontal axes.

one cue was sighted, has an estimated abundance of 200 ($CV=111.1\%$) from the stratified analysis, compared with the spatial modelling estimate of 1,150 ($CV=40.4\%$ from the jackknife or 34.8% from the parametric bootstrap). The stratified estimate, albeit with a large CV, stands out as being very low compared with both the stratified estimates and the estimates from the spatial model from the adjacent blocks 2BC and 4C.

The low number of sightings in this survey generally resulted in low estimator precision from the stratified analysis. The spatial model achieved substantially higher precision for the total estimate of abundance, both from the jackknife and the parametric bootstrap, the parametric bootstrap also giving improved precision for the stratum estimates. The jackknife CVs were lower than those from the stratified analysis in blocks with few sightings, whilst in blocks 4C and 5C, the higher number of cues improved the performance of the analytic estimator and in both cases the CVs were 8.6% lower than the jackknife CVs.

As for the Antarctic minke whale example, the greatly improved precision from the spatial model when the parametric bootstrap is used may be due to failing to incorporate fully the correlation structure present in the original data. The estimated dispersion parameter

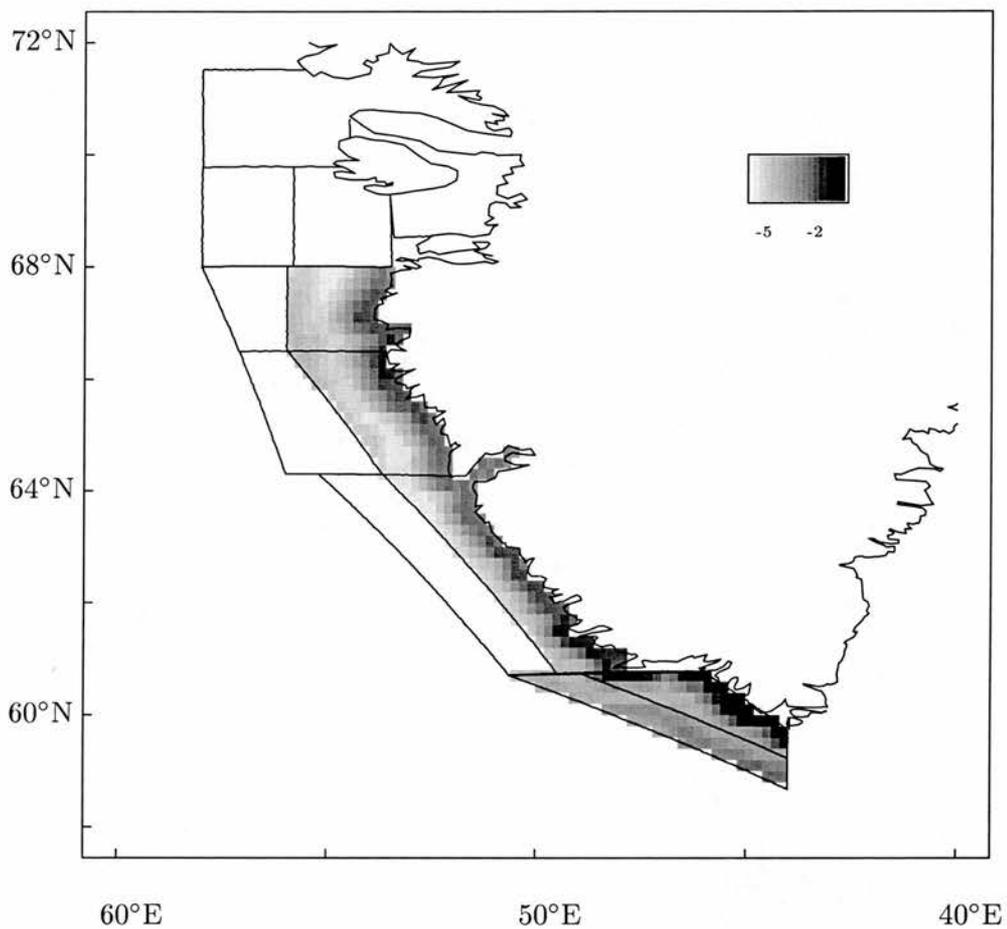


Figure 8: The logarithm of the predicted density of minke whale schools in the surveyed region, predicted from a model based on cubic smoothing splines of distance from the coastline and latitude, and a linear longitude term. To avoid extrapolating beyond the range of the data, the predicted surface is only plotted for survey blocks in which sightings were made.

from the model for the original data was 6.13, compared with an estimate of 4.32 for the mean of the dispersion parameter estimates from the bootstrap samples. A Monte Carlo test carried out as described in section 6.1 indicated that this difference was significant (p -value=0.02), so the CVs from the parametric bootstrap are indeed likely to be negatively biased.

Block	Number of Cues	Stratified Analysis		Spatial Model		
		\hat{N}	%CV	\hat{N}	%CV _J	%CV _{PB}
2BC	7	1,540	85.2	1,070	76.5	43.5
3C	1	200	111.1	1,150	40.4	34.8
4C	19	1,180	47.6	2,030	56.2	31.3
5C	27	3,160	52.0	2,320	60.6	34.3
5O	4	1,320	95.7	640	84.8	50.3
TOTAL	58	7,400	41.1	7,210	33.20	28.60

Table 3: Comparison of estimates of abundance of minke whale schools (\hat{N}) from a conventional stratified analysis (Hedley *et al.*, 1997b) and the spatial modelling approach. For the spatial model, the CVs were estimated using the jackknife (denoted %CV_J) and the parametric bootstrap (%CV_{PB}).

7 Discussion

Although the use of generalized linear or generalized additive models for estimating the spatial distribution of wildlife distribution is already an established approach when data are available at sample points throughout a survey grid, they have not been widely used to model data from line transect surveys. In this chapter, new methodology (recently published in Hedley *et al.* (1999)) for estimating the spatial distribution of wildlife from strip transect and line transect data was presented. It was also shown how the methodology could be adapted for data collected on a cue-counting survey.

The basis of the method was to subdivide transects into small segments and model the number of groups in each segment. This approach is straightforward in the case of strip transects, but when the detectability of groups varies according to distance from the ob-

server, and possibly other covariates, then the method requires the probability of detection of groups to be incorporated into the modelling framework. In the simple case, where estimated probability of detection is the same for every detected group, it was suggested that this estimated probability should enter the model via the offset term. Whilst it is non-standard to have an estimated quantity as an explanatory term in the model, this presents no modelling difficulties in practice. The fact that an explanatory variable has itself been estimated should be accounted for in the variance estimation procedure; both the jackknife and parametric bootstrap described in this chapter do this.

These resampling techniques were proposed as candidate methods to estimate the variance of density and abundance estimates from the spatial model. Analytic estimators would be inappropriate because the division of the transects into segments causes non-independence in the observations. The performance of these techniques requires further scrutiny before the estimates of precision can be used with confidence. On the face of it, the jackknife results are moderately encouraging but there may be several reasons in practice why one might choose not to use this estimator. Firstly, the number of resamples is constrained by the number of transects; it is most appropriate for surveys with a large number of transects. Additionally, because these transects should be of approximately equal length, the number of surveys which might produce data suitable for the jackknife will be limited. The statistical properties of the jackknife estimator are such that it operates correctly only if the estimated statistic ($\hat{\theta}$, say) has a locally linear quality (Miller, 1974). Intuitively this means that $\hat{\theta}$ should exhibit ‘smoothness’, whereby small changes in the data cause only small changes in the value of the estimated statistic (Efron and Tibshirani, 1993). However the jackknife estimator of abundance ought to satisfy this smoothness criterion, and it is probably more the nature of the observations themselves that suggest that the jackknife will only rarely be appropriate for estimating variance from the spatial models described here. In particular, Miller (1974) notes that it is uncommon for the jackknife to be successful for time series applications. The line transect data, once transformed into counts per segment along the lines, may be thought of as a series of clusters of observations which are serially correlated. Such data are clearly analogous to time series data.

The improvements in precision from the parametric bootstrap in the two applications presented in section 6 are very promising, but the Monte Carlo test comparing the dispersion parameters provided some evidence that the true variance was being underestimated.

This is likely to have been caused by a failure to reproduce all of the spatial correlation between neighbouring segments. In recently presented work, Bravington (2000a) pursues an approach analogous to those used in time series applications in order to incorporate autocorrelation directly. An alternative approach might be to include an autocovariate term in the model as in Augustin *et al.* (1998), and use Markov Chain Monte Carlo (MCMC) methods to enable prediction from the fitted model.

A better comparison between stratified line transect methods and the methodology presented in this chapter would be gained from using simulated data to examine possible biases and to evaluate conditions under which the spatial modelling approach might be expected to perform favourably. In the next chapter we introduce an alternative spatial modelling formulation which is based on the distance between successive sightings.

Chapter 4

An Alternative Perspective: Using Intervals Between Sightings

1 Introduction

Although the count data model has been applied to several quite different data sets yielding informative results (for example, Hedley *et al.* (1997b); Clarke *et al.* (2000)), the alternative framework presented in this chapter was originally motivated by an attempt to develop methodology with more theoretical appeal. In particular, the segmentation of the transect lines is rather *ad hoc* in nature, and yields artificially many observations. Therefore, an alternative approach is suggested in this chapter which still may be implemented within a GLM or GAM framework, but instead of counts, the response is based on intervals between successive sightings. In the simple case of equal probability of detection for all sightings, the interval is the distance surveyed between sightings (hereafter called the ‘waiting distance’). If probability of detection varies spatially according to environmental conditions, the response becomes the ‘waiting area’, defined as the waiting distance multiplied by the appropriate estimated effective strip width.

2 Constant expected encounter rate model

In this section, a model is developed under the assumption that the density of groups and the expected encounter rate is constant in the interval between one detection and the next. In reality, this is clearly an implausible model, but it serves as a starting point for a more appropriate model which must be fitted iteratively (and this is described in section 3).

2.1 Derivation

Suppose we have a vector of spatial covariates, \mathbf{z} , whose values are observed throughout a survey region on a fine grid. We begin with a conceptual model in which these covariates, the density of groups and the expected encounter rate all remain constant as the observer travels from one detection to the next, but may change when a detection occurs. Let ξ be a vector with components (x, y) denoting the spatial location of a detection in the survey region, relative to orthogonal axes x and y , where (without loss of generality) the x -axis is defined to be along the trackline. Hence, the y -axis denotes the perpendicular distance of the detection from the trackline. The locations of groups (detected and undetected) are assumed to follow an inhomogeneous Poisson process, with rate parameter, $D(x, y)$, varying with location in the survey region. For this conceptual model, the process equates to a sequence of homogeneous Poisson processes, where the rate of each process may differ, but is constant between any two detections.

Suppose that detections occur at $(x_i, 0)$ and $(x_{i+1}, 0)$, where $x_{i+1} = x_i + l_{i+1}$ and l_{i+1} is the along-trackline distance between the two detections. For this conceptual model, we assume that the density of groups between the two detections is a constant equal to the density, $D(x_{i+1}, 0)$, at $(x_{i+1}, 0)$. For simplicity, we will also assume that the estimated effective strip half-width, $\hat{\mu}$, is constant throughout the survey region.

Under these conditions, the expected number of detections within the strip of length l_{i+1} and width $2\hat{\mu}$ is $2\hat{\mu}l_{i+1}D(x_{i+1}, 0)$.

Now let M_i be a random variable denoting the length of trackline surveyed before a detection is made, starting from $(x_i, 0)$. The cumulative distribution function (CDF) of M_i evaluated at $(x_{i+1}, 0)$ is:

$$F_{M_i}(l_{i+1}|x_i) = P(M_i \leq l_{i+1}|x_i) = 1 - P(M_i > l_{i+1}|x_i).$$

Since $M_i > l_{i+1}$ if and only if there were no detections in the strip of width $2\hat{\mu}$, then

$$\begin{aligned} F_{M_i}(l_{i+1}|x_i) &= 1 - P(\text{no detections in the strip}) \\ &= 1 - \frac{[2\hat{\mu}l_{i+1}D(x_{i+1}, 0)]^0 \exp\{-2\hat{\mu}l_{i+1}D(x_{i+1}, 0)\}}{0!} \\ &= 1 - \exp\{-2\hat{\mu}l_{i+1}D(x_{i+1}, 0)\}, \end{aligned} \quad (1)$$

ie. the exponential distribution, $\text{Exp}(\phi)$, with $\phi = 2\hat{\mu}D(x_{i+1}, 0)$. The mean of this distribution is $1/\phi$ and its reciprocal, ϕ , is the intensity of a homogeneous Poisson process in the area of length l_{i+1} and width $2\hat{\mu}$ between $(x_i, 0)$ and $(x_{i+1}, 0)$. The observed density at $D(x_{i+1}, 0)$ is $1/(2\hat{\mu}l_{i+1})$.

2.2 A spatial model based on waiting distances

It follows from the previous section that an appropriate response variable for a spatial model is the waiting distance. The form of the GLM,

$$g[E(l_i)] = \theta_0 + \sum_k \theta_k z_{ik}, \quad i = 1, \dots, n, \quad (2)$$

is similar to that of the GAM,

$$g[E(l_i)] = \theta_0 + \sum_k f_k(z_{ik}), \quad i = 1, \dots, n, \quad (3)$$

where the link function g is a monotonic differentiable function (the logarithmic link may often be suitable, as it ensures positive values of the mean response), k is the number of spatial covariates included in the model, and n is the number of detected groups.

If it is assumed that the l_i , $i = 1, \dots, n$, are exponentially distributed, then the gamma distribution, with dispersion parameter set to unity, is the appropriate error distribution

(since this is the exponential distribution). In practice, however, the data are likely to be overdispersed. The gamma distribution is still appropriate, but the dispersion parameter must be estimated.

After fitting a model to the observed waiting distances, the predicted waiting distance (and hence predicted density) can be evaluated for the grid of spatial covariates throughout the survey region.

This formulation can only be fitted as a GLM/GAM because of the assumption of constant density between detections. Standard GLM/GAM software cannot be used to fit an inhomogeneous Poisson process, where the rate is a function of spatial location.

3 Variable expected encounter rate model

Constrained by the desire to use standard statistical software, in this section we build on the conceptual model introduced above, and suggest an iterative method that accommodates variation in density, in the spatial covariates and in the expected encounter rate between successive detections.

The iterative procedure is such that the observed waiting distances are adjusted to the distances that would have occurred if the underlying (inhomogeneous) Poisson process was in fact homogeneous and the density between detections was constant.

The first step of the iteration is to fit a spatial model to the observed waiting distances as described in section 2.2. The adjusted waiting distances, \tilde{l}_i , $i = 1, \dots, n$, satisfy:

$$1 - \exp \left\{ -2\hat{\mu} \int_{x_{i-1}}^{x_{i-1}+l_i} \hat{D}(x, 0) dx \right\} = 1 - \exp \{ -2\hat{\mu} \tilde{l}_i \hat{D}(x_i, 0) \}.$$

Taking logarithms, this simplifies to:

$$\int_{x_{i-1}}^{x_{i-1}+l_i} \hat{D}(x, 0) dx = \tilde{l}_i \hat{D}(x_i, 0).$$

Therefore, the adjusted distances are calculated as

$$\tilde{l}_i = \frac{\int_{x_{i-1}}^{x_{i-1}+l_i} \hat{D}(x, 0) dx}{\hat{D}(x_i, 0)}, \quad (4)$$

ie. the area under the predicted density surface between detections at $(x_{i-1}, 0)$ and $(x_i, 0)$ is equated to the area of the rectangle of width \tilde{l}_i and height $\hat{D}(x_i, 0)$, as illustrated in Figure 1.

Given the fitted model, the integral $\int_{x_{i-1}}^{x_{i-1}+l_i} \hat{D}(x, 0) dx$ may be evaluated numerically. The denominator, $\hat{D}(x_i, 0)$, is calculated as the reciprocal of the i^{th} fitted value multiplied by the estimated effective strip width, $2\hat{\mu}$.

The model is then refitted to the adjusted distances to give an estimate of the density surface $D(x, y)$ with smaller bias. This process is repeated, each time adjusting the waiting distances, until convergence is reached.

4 Variance estimation

In chapter 3, section 5.1, we outlined three possible implications of modelling dependent observations within a GLM/GAM framework. Two of these related to likelihood-based model selection criteria; the other noted that analytical estimates of variance would be biased. Except for the case when objects are distributed according to a homogeneous Poisson process throughout the survey region (ie. they are randomly distributed in the region — a situation which rarely occurs in practical applications), the waiting distances will be dependent observations. Therefore, as for the count data models described in chapter 3, an appropriate resampling method should be used to obtain robust estimates of variance.

For the reasons discussed in chapter 3, the non-parametric bootstrap will generally be inappropriate, and this method is not considered further here. Instead, attention is again focussed on the jackknife and the parametric bootstrap.

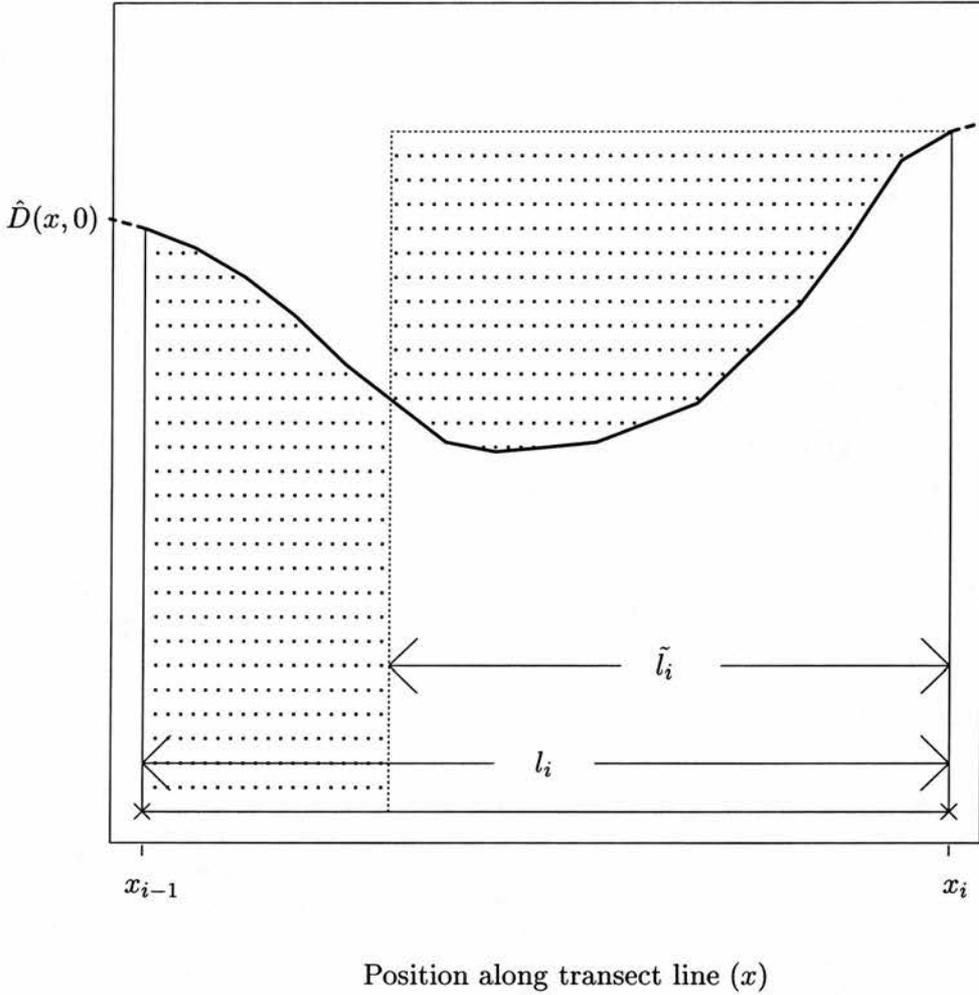


Figure 1: Density on the trackline, $\hat{D}(x, 0)$, with detections at x_{i-1} and x_i . The distance between detections is shown as l_i ; the adjusted waiting distance is shown as \tilde{l}_i . \tilde{l}_i is such that the area of the rectangle of width \tilde{l}_i and height $\hat{D}(x_i, 0)$ equals the area under $\hat{D}(x, 0)$ between x_{i-1} and x_i (ie. so that the dotted regions are equal in area).

4.1 The jackknife

As in chapter 3, the sampling units (which are assumed to be independent) are defined as the transect lines (or alternatively days of survey effort). The jackknife samples are then obtained by systematically deleting one transect (or day) at a time from the observed data. Suppose that the sampling units are taken to be the lines and there are τ of these in the original data, then each of the τ jackknife samples are analysed as if they were the real data.

This first involves re-estimation of the effective strip width, to obtain estimates $\hat{\mu}^{(-t)}$, where the $(-t)$ superscript indicates that the estimate is calculated with the t^{th} transect removed. The second step involves, for each resample, recalculation of the waiting distance affected by the removal of a transect. This will usually entail calculating the waiting distance from the last detection on transect $(t - 1)$ to the first detection on transect $(t + 1)$. The corresponding intervals for numerically integrating between these two detections in order to carry out the iterative procedure must also be redefined. Neglecting model selection uncertainty, the spatial model fitted to the original data is then fitted to the jackknife samples, the iterative procedure being carried out for each resample. (It was noted in chapter 3 how to incorporate model selection uncertainty, if required.)

Estimates of abundance from the jackknife samples, $\hat{N}^{(-t)}$, $t = 1, \dots, \tau$, may be calculated by integrating under the density surface predicted from the final iteration of the respective model fits.

The jackknife estimator of the variance of the abundance estimate (from the original data), \hat{N} , is calculated by first constructing τ pseudovalues, $\tilde{N}^{(-t)}$, $t = 1, \dots, \tau$. The variance of \hat{N} is then given by:

$$\text{var}(\hat{N}) = \frac{\sum_{t=1}^{\tau} (\tilde{N}^{(-t)} - \tilde{N}^{(\cdot)})^2}{\tau(\tau - 1)}, \quad (5)$$

where $\tilde{N}^{(\cdot)}$ is the mean of the abundance from the pseudovalues.

4.2 The parametric bootstrap

Recall that for the parametric bootstrap, instead of resampling from the observations, resamples are taken from the model fitted to the original data. An algorithm for obtaining parametric bootstrap resamples in the context of spatial models from count data was given in chapter 3, section 5.4. For the waiting distance models, steps 1-4 of this algorithm are the same. Step 5 becomes

5. The along-trackline location of the groups is given by the accepted $\text{Unif}(0, L)$ deviates. The corresponding spatial covariates at these locations must be obtained.

Neglecting model selection uncertainty, the spatial model fitted to the original data is then fitted to the parametric bootstrap resamples, the iterative procedure being carried out for each resample. An estimate of abundance in the survey region from each of the resamples may be obtained by numerically integrating beneath the appropriate region of the predicted density surface. The variance of \hat{N} comprises two components: the first is due to estimation of the effective strip width, and the second is due to the spatial modelling. The sample variance of the abundance estimates from the parametric bootstrap resamples is an estimate of the component of variance of \hat{N} due to the spatial modelling. The delta method (Seber, 1982) may be used to combine this component of the variance with the component of the variance due to estimation of the effective strip width, and thus obtain an estimate of the overall variance of \hat{N} .

5 Application of the methodology

5.1 Antarctic minke whales

In this example, we apply the interval data spatial modelling methodology to Independent Observer (IO) mode minke whale data from the 1992-3 IWC/IDCR Antarctic Survey in Area III. These data were also used in the first example given in the previous chapter (section 6.1).

As in chapter 3, the analysis follows that of Borchers and Cameron (1995). Thus effective

strip widths were estimated separately for the northern strata (WN and EN) and the southern strata (WS and ES), and are shown again in Table 1 below.

Pooled strata	\hat{p}	%CV	$\hat{\mu}$ (n.miles)	%CV
WN and EN	0.742	7.61	1.112	7.61
WS and ES	0.360	15.42	0.540	15.42

Table 1: Stratum estimates of detection probability, \hat{p} , of groups within the strip of half-width 1.5 n.miles, and effective strip half width, $\hat{\mu}$, with coefficients of variation (calculated using DISTANCE software).

Waiting distances (l_i , $i = 1, \dots, n$) were calculated as the along-trackline distances between successive detections, irrespective of whether or not any two detections were on the same transect line. For each of the two vessels, there was a period of on-effort searching after that vessel's last detection until the end of the trackline. The distances corresponding to this effort were added to the (left-censored) waiting distance of the first detection from each vessel. Since the estimated effective strip width is not constant throughout the survey region, the methodology described in section 3 was slightly modified. Instead of modelling the observed waiting distances, the response was the estimated waiting area, \hat{r} (say), calculated as $\hat{r}_i = 2l_i\hat{\mu}$, $i = 1, \dots, n$, where $\hat{\mu} = 1.112$ if the i^{th} detection is in one of the northern strata, and $\hat{\mu} = 0.540$ if it is one of the southern strata.

A generalized additive model with a logarithmic link function was fitted to the estimated waiting areas, assuming a gamma error distribution. As for the count data model, three spatial covariates were available: distance from the ice edge (*icedge*), latitude (*lat*) and longitude (*lon*). These were considered for inclusion in the model as cubic smoothing splines with either 8, 4 or 2 degrees of freedom, or as linear terms. Stepwise automated model selection, based on the AIC (adjusted for overdispersion), led to the following final model:

$$E(\hat{r}_i) = \exp[\theta_0 + s(\text{lat}_i, 8) + s(\text{lon}_i, 2)] , \quad i = 1, \dots, n, \quad (6)$$

where n is the total number of detections.

The failure to select the term representing the distance from the ice edge was counter to our *a priori* expectations, but in this case, the high flexibility of the smoothed term in latitude, coupled with the strongly smoothed longitudinal term, was sufficient to model the spatial variation in density. The nonlinear form of the dependence of the estimated waiting areas on these two covariates is shown in Figure 2. Note that the increasing trend with decreasing latitude seen in the smoothed function of this covariate represents a decreasing trend in density (because expected densities are given by the reciprocal of the expected waiting areas). The predicted density surface of minke whale schools is shown in Figure 3.

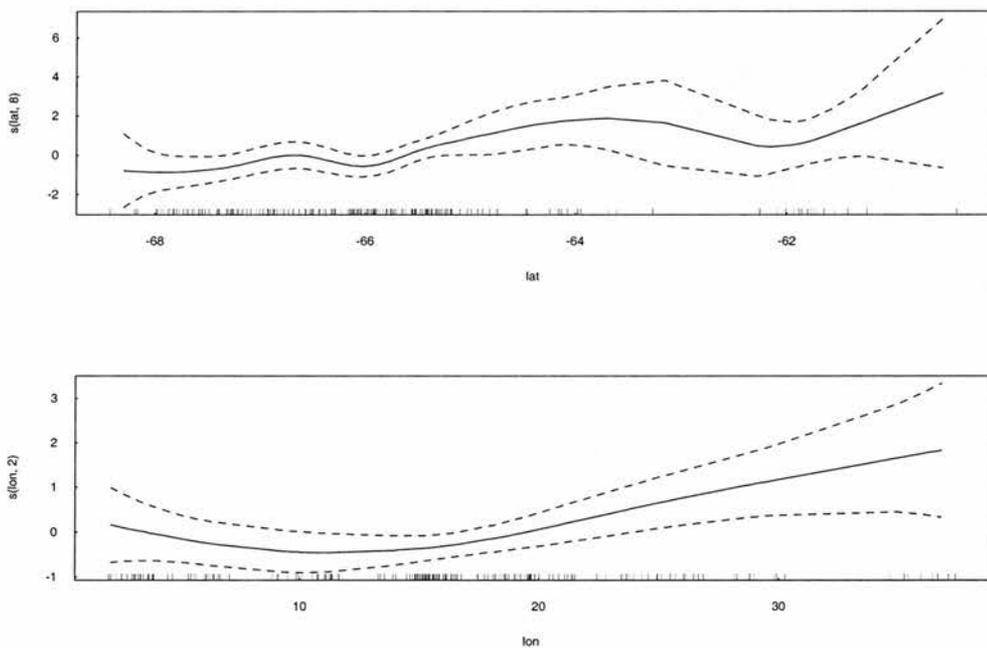


Figure 2: Shapes of the functional forms for the smoothed covariates, latitude and longitude. Zero on the vertical axes corresponds to no effect of the covariate on the estimated response (which for this model is waiting area = [pod density]⁻¹). The locations of the observations are plotted as small ticks along the horizontal axes.

Estimates of abundance from the waiting areas model are compared with those from the stratified analysis of Borchers and Cameron (1995), and with those from the count data model from chapter 3, in Table 2. For the two spatial models, coefficients of variation were estimated using the jackknife and the parametric bootstrap.

Since the primary purpose of this example was to demonstrate how the methodology might be applied (rather than to fit the ‘best’ possible model), the conclusions that can

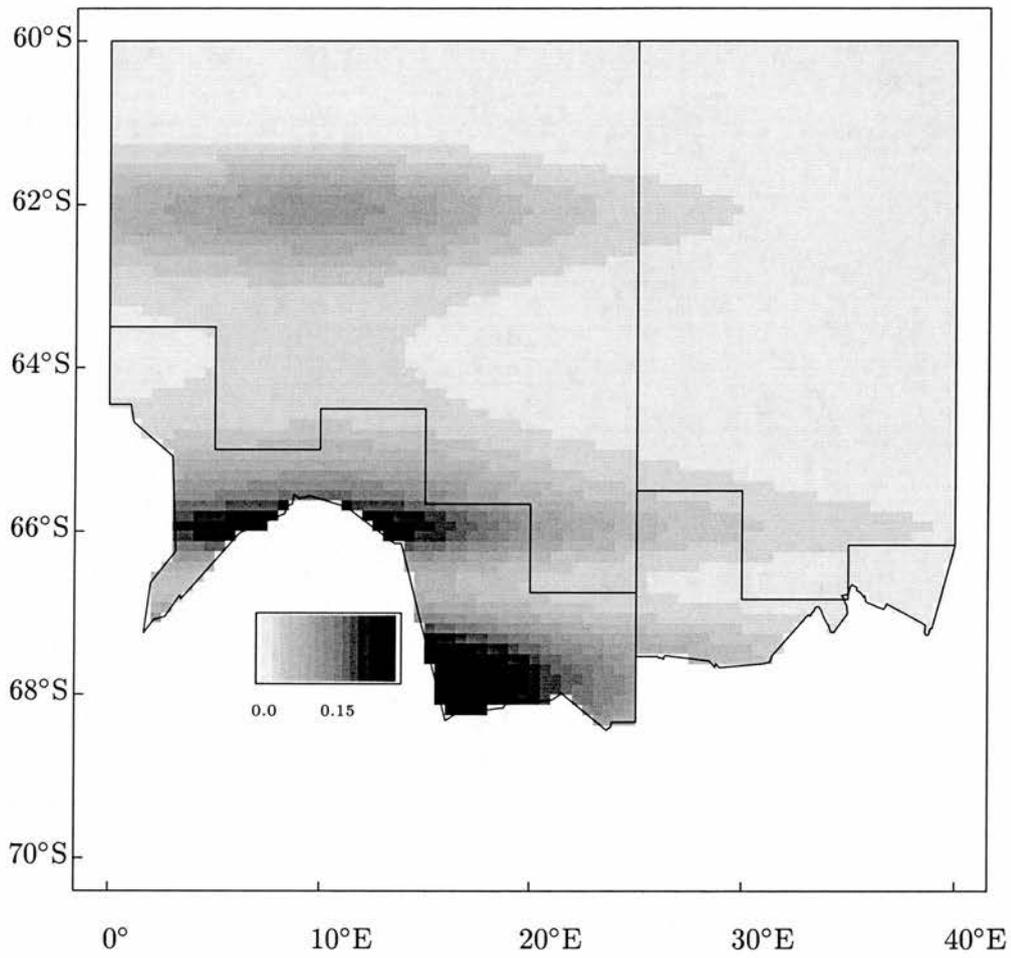


Figure 3: Density of minke whale schools in the surveyed region, predicted from a model based on cubic smoothing splines of latitude and longitude.

Stratum	Stratified Analysis		Count Data Model			Waiting Areas Model		
	\hat{N}	%CV	\hat{N}	%CV _J	%CV _{PB}	\hat{N}	%CV _J	%CV _{PB}
WN	4,810	40.1	4,620	40.4	19.9	6,386	56.9	30.6
EN	1,460	49.5	820	44.4	28.9	1,058	43.4	31.3
WS	7,410	25.1	8,420	30.6	17.6	6,895	45.0	23.0
ES	640	44.3	880	43.3	25.1	686	69.6	29.1
TOTAL	14,320	23.0	14,740	25.5	16.0	15,025	39.2	21.2

Table 2: Comparison of estimates of abundance of minke whale schools (\hat{N}) from a conventional stratified analysis (Borchers and Cameron, 1995), the count data model (chapter 3) and the waiting areas model described in this section. For the two spatial models, the CVs were estimated using the jackknife (denoted %CV_J) and the parametric bootstrap (%CV_{PB}).

be drawn from comparing the results between the three approaches should only be considered preliminary. In particular, the spatial covariates and the number of smoothing parameters for the spatial models were selected using an automated, AIC-based model selection procedure, which we know will tend towards overfitting when the assumption of independent observations is invalid. Nevertheless, the point estimates of total abundance from the three approaches are remarkably similar. Whilst the estimate of abundance in the WN stratum from the waiting areas model is higher than the corresponding estimates from the other two methods, the differences are unremarkable considering the precision of the estimates. However, it may be that in this case the model is a poor fit to the data, and influenced by the eleven sightings clustered on the westernmost transect of the WN stratum, fails to capture the more expected scenario that density decreases with distance from the ice edge.

As with the count data model, for the estimates of abundance by stratum (and most markedly for the three strata with fewest sightings), the CVs for the estimates from the waiting areas model calculated using the parametric bootstrap suggest considerable improvement in precision compared to the CVs from the stratified analysis. A marginal gain in precision was also obtained for the total abundance estimate. The CVs estimated by the jackknife are rather worse. They compare poorly in all respects — the parametric

bootstrap estimates from the same model, the jackknife estimates from the count data model and the stratified analysis estimates are all substantially lower, with only one exception (the jackknife estimate for the EN stratum is lower than that of the stratified analysis). Possible explanations for the failure of the jackknife were presented in section 7 of the previous chapter.

The improvement in precision gained by the spatial model when the parametric bootstrap was used for variance estimation is encouraging, but since it is known that parsimonious models for the spatial variation in density are unlikely to be sufficiently flexible to incorporate fully the serial correlation between successive observations, the estimates are likely to be negatively biased. The amount of overdispersion present in the resamples compared to that in the original data should, to some extent, reflect whether the resamples exhibit as much serial correlation as seen in the original data. The estimated dispersion parameter from modelling the original data was 3.22; the mean of the dispersion parameter estimates from the resamples was somewhat lower, at only 1.92. A Monte Carlo test indicated that the difference was significant (p -value=0.04), leading to the conclusion that there was likely to be more variability in encounter rate in the original data than in the resamples, and evidently confirming that the parametric bootstrap estimates in Table 2 are negatively biased.

6 Further research

6.1 A convergence problem

It might appear strange to note this problem *after* the methodology has been applied to a real data set, but since it only came to light during the implementation, it seemed appropriate to present it here.

In section 3, the following statement was made:

This process is repeated, each time adjusting the waiting distances, *until convergence is reached*.

It transpired that the latter part of this sentence was a bit too optimistic!

In the example, ‘convergence’ was considered to have been achieved when the sum of the absolute differences between the adjusted waiting areas from successive iterations was less than 5 n.mile². (This level of precision is probably excessive, since the estimates generally changed little between iterations, once the first one or two adjustments had been made.) However, it was apparent that for some data sets, convergence did not occur. Visual inspection of the data revealed no obvious differences between those data sets for which the iterative modelling successfully converged and those for which it did not. Upon closer examination of equation 4,

$$\tilde{l}_i = \frac{\int_{x_{i-1}}^{x_{i-1}+l_i} \hat{D}(x, 0) dx}{\hat{D}(x_i, 0)},$$

it is evident that if any of the n fitted values of density, $\hat{D}(x_i, 0)$ are equal or nearly equal to zero, then the corresponding adjusted waiting distance/area tends to infinity.

In the example, in order to ensure convergence, both for the original data and for the resamples generated to estimate variance, a constraint was applied to the adjusted waiting areas. This was such that if an adjusted waiting area exceeded some maximum value, then it was set to that value. The maximum value used was the total waiting area, i.e. the total survey effort multiplied by the appropriate estimate of effective strip width. An alternative maximum value has been suggested: the reciprocal of the minimum possible density with at least one detection, i.e. $[1/A]^{-1}$, where A is the total area of the survey region. This constraint has not been tried, but it should work equally well.

We do not claim to have solved the convergence problem in general with this rather *ad hoc* fix, although it was adequate for our purposes in the example. Clearly, the convergence properties of the iterative procedure still require some further investigation.

6.2 Discussion

The spatial modelling methods presented in this chapter and in chapter 3 represent a promising improvement over traditional line transect estimation methods in several respects. Potentially, the most useful applications of the methodology might be:

1. To provide a statistically sound means for estimating abundance at any spatial res-

olution, for example, by Small Management Area (IWC, 1999), with relatively high precision.

2. To provide higher precision for abundance estimation in the whole survey area than stratified estimation methods.
3. To provide a powerful tool for relating animal distribution and abundance to spatial and other explanatory variables.

Although the results from the examples provide some empirical evidence to support the first two statements, the Monte Carlo tests examining the dispersion parameters suggest that the parametric bootstrap — our favoured choice of variance estimator — is negatively biased. The extent of the bias needs to be established before the associated variance estimator can be used with confidence, e.g. for management purposes. An alternative and probably better approach would be to incorporate the correlation structure directly. Gotway and Stroup (1997) present a framework for combining generalized linear models and quasi-likelihood with geostatistical methods, defining a general variance-covariance matrix that accounts for autocorrelation. They initially fit a GLM which models spatial trend, and then estimate the autocorrelation structure from the semi-variogram of the residuals. The estimated correlation matrix is then inserted into the variance-covariance matrix when fitting the final GLM, using quasi-likelihood. All three examples presented are based on discrete data, but in principle the methodology could also be applied to continuous data. It would be an interesting and possibly fruitful avenue for further research in exploring this approach for use with the spatial models presented in this and the previous chapter, and in particular to extend the GLM framework in Gotway and Stroup (1997) to a GAM framework. As noted in chapter 3, a primary objective of time series analyses is modelling the autocorrelation structure in the data, therefore by modifying the techniques commonly used in such analyses, they might become applicable in the spatial context. Such modifications have recently been investigated (Bravington, 2000a).

In this thesis, we have not been concerned with exploring the potential of the spatial models to relate animal distribution and abundance to explanatory variables, at least not ecologically meaningful ones. The primary development has been in constructing a framework in which such modelling can take place, and the examples have used locational covariates (latitude, longitude and distance from the ice edge or coastline) which only serve

as proxies for other covariates which might genuinely be expected to influence habitat selection, such as food resource, sea surface temperature and upwellings. Augustin (1999) remarks that spatial modelling might be carried out for at least two fundamentally different purposes: spatial modelling for estimating global characteristics and spatial modelling for the description of the underlying process. Abundance estimation falls into the first category, and Augustin (1999) notes that estimation is fairly robust to overfitting of the model for such a purpose. However, it seems critical that model selection criteria with a substantive theoretical basis are developed for GAMs, if we are to use them with confidence to infer relationships between animal distribution and habitat covariates. Some progress in this regard has been made by Wood (2000), for cubic smoothing splines.

A common research objective is to relate not only spatial variation in animal distribution and abundance with environmental variables, but also to examine the corresponding temporal variability. The spatial modelling methodology provides a useful tool for addressing this objective. With sufficient coverage over a time period, it would be quite straightforward to incorporate a temporal component to assess changes in spatial distribution with time. Whilst this may result in useful inferences over time periods in which biological processes are not expected to affect substantially the population under study (for whales, for example, a year; for mice, less than a month), more useful inferences about the temporal variability of the population are likely to result from methods which integrate process models and survey data (e.g. Sullivan (1992); Fewster (1999b); Trenkel *et al.*(2000)). Spatial models like those presented in this and the previous chapter are potentially an important component of such integrated models.

As mentioned at the beginning of this chapter, the interval data methodology was originally developed to try to obtain spatial models with a more theoretical foundation than those developed in chapter 3. In fact, we have seen that many of the pitfalls of the count models, e.g. dependence of the observations of the response, negative bias in the variance estimates, the need to estimate a parameter outside the spatial modelling (p_{ij} or effective strip width), have not been overcome by modelling the intervals between sightings.

In this chapter, we have restricted attention to the development of models which could be fitted using standard statistical GLM/GAM software. In the following chapter, we consider extensions to this approach and present a more general framework for spatial modelling from line transect data.

Chapter 5

Point Process Modelling of Line Transect Data

1 Introduction

In the previous two chapters, two modelling formulations were presented for estimating the spatial distribution of wildlife from line transect data. The models were developed in either a GLM or a GAM framework, allowing implementation within widely used statistical software, such as S-PLUS. In this chapter, a modelling structure is developed that is based on the formulation of a general likelihood framework to estimate the spatial variation in density from line transects. It is suggested that point processes (see for example, Cox and Isham (1980)) provide an appropriate framework for such modelling.

Point process formulations of line transects estimators were considered by Stoyan (1982) and Högmänder (1991; 1995). It was Stoyan (1982) who originally reformulated the conventional line transect estimator of group density

$$\hat{D} = \frac{n}{2L\hat{\mu}} \quad (1)$$

in terms of the intensity of a stationary marked point process, where the points were the locations of the animal groups, and individual points were *marked* by some probability representing the detectability of each point, based on its sighting distance. Stoyan (1982)

justifies the assumption of stationarity by stating that a usual assumption in conventional line transect estimation is that the animals are uniformly distributed throughout the region of interest. In fact, conventional line transect estimation does not require this assumption provided that the transect lines are placed at random with respect to the locations of animal groups. The stationarity assumption simplifies the point process formulation substantially but is not realistic for real applications. Both Stoyan (1982) and Högmänder (1991; 1995) were primarily concerned with demonstrating the equivalence of the stationary point process formulation and the conventional estimator in the case when some ‘average’ intensity is the parameter of interest to be estimated (corresponding in line transect parlance to the average density of animal groups). In contrast, in this chapter, we attempt to address the problem of estimating the density of animal groups when this is modelled as a function of spatial location.

1.1 Some definitions and notation

A *point process*, Φ , is a model for identical objects, called *points* (denote these by X_i for now), which are distributed in some space. In this chapter, attention will be restricted to planar point processes, therefore the points will be distributed in \mathbb{R}^2 . Furthermore, we shall only be concerned with *simple* point processes, which means that all points are distinct (ie. multiple points arising from objects being at exactly the same location will not be considered).

Thinning of a point process is an operation which deletes points from the process Φ according to some rule. The simplest thinning is *p-thinning*, where $p \in (0, 1)$: each point of Φ is deleted with probability $1 - p$ independently of its location and of deletions or retentions of the other points of Φ . A non-stationary generalization of *p-thinning* is obtained by defining a retention probability function $p(x)$, where $0 \leq p(x) \leq 1$. Thus, the thinning depends on the location of the point; a point $x \in \Phi$ is deleted with probability $1 - p(x)$, independently of the deletions or retentions of the other points. This operation is called *p(x)-thinning*.

A point process is made into a *marked* point process by attaching a characteristic (the *mark*) to each point of the process. The mark is itself a random variable. If Φ is a point process and M_i are i.i.d. random elements of the set of possible marks (the *mark space*)

and also independent of Φ , then the marked point process, Ψ , is said to be obtained from Φ by *independent marking*. If the marks, M_i , are conditionally independent given Φ , but each mark M_i depends on the underlying point X_i , Ψ is said to be obtained from Φ by *position-dependent marking*.

A marked point process could also be thinned with respect to its marks. This kind of thinning is either dependent or independent according to whether the marking is dependent or independent of (the locations of) the point process.

1.2 Putting these definitions into a line transect context

In the line (or strip) transect context, the locations of objects (say, animal groups), define a point process in the survey region, which is a bounded subset of \mathbb{R}^2 . Since animals typically favour certain habitat or environmental conditions, their distributions are generally clustered and it is therefore unreasonable to assume that their locations would be described well by a stationary process. It is likely that the groups' locations would be better modelled by a non-stationary process, such as the inhomogeneous Poisson process, where the rate of the process (effectively the expected number of animal groups per unit area) is allowed to vary as a function of spatial location.

Following the notation of chapter 4, let the rate of an inhomogeneous Poisson process describing the location of animal groups be denoted $D(x, y)$, where x denotes position along the direction of the transect line and y denotes position in an orthogonal direction to the transect line (therefore $|y|$ is 'perpendicular distance' in the conventional line transect sense).

An important consequence of modelling the groups' locations using the inhomogeneous Poisson process is that if A_1, \dots, A_k are arbitrary disjoint sets within the survey area, A , then the numbers of animal groups in each set, $N(A_i)$, $i = 1, \dots, k$, are still independent Poisson variables, with the following expected values:

$$E[N(A_i)] = \int_{A_i} D(x, y) dx dy, \quad i = 1, \dots, k. \quad (2)$$

The surveyed 'strips' (of area $2L_i w$) from a line transect survey, where L_i is the length of

the i^{th} transect and w is the perpendicular truncation distance, may be viewed as disjoint sets within the survey region. Consequently, the number of groups within each surveyed strip are still Poisson variables and their locations within the strip can be modelled using an inhomogeneous Poisson process.

The statistical problem is to estimate the spatial density surface, $D(x, y)$, throughout the survey region, A , from the observed data — the locations of detections with respect to the transect lines. Therefore, now consider the *detected* animal groups from a line transect survey. Conventional line transect estimation requires estimation of a detection function, $g(y)$, $0 \leq g(y) \leq 1$, which is typically assumed to be symmetric about the line so that $g(y) = g(-y)$. The sign of y is therefore ignored. It is also usually assumed that animals on the line are detected with certainty so that $g(0) = 1$, although methods exist to accommodate uncertain detection on the trackline (as discussed in chapter 2). In addition to modelling the probability of detection of groups depending on their perpendicular distance from the line, variability in their detectability at different locations along the line can be modelled using a bivariate detection function $g(x, y)$. By analogy with the conventional definition of $g(y)$ (as given, for example, in Buckland *et al.* (1993a), page 9), $g(x, y)$ may therefore be defined as the probability of detecting a group, given that (the centre of) the group is at location (x, y) . This definition of the detection function readily accommodates ‘weather’ covariates such as sea state, where the covariate, z_1 say, may be recorded at locations $(x, 0)$ along the transect line, or possibly at the location of the group itself. Thus the detection function may take a form such as $g(z_1(x), y)$ or $g(x, y, z_1)$ respectively. However, the framework does not readily incorporate sighting-specific covariates such as group size. For clarity, the notation $g(x, y)$ will be used to represent spatial variability in the detection function. Implicitly, this may be taken to represent a detection function depending on some appropriate covariates, denoted by the vector \mathbf{z} ; hence $g(x, y)$ may be read equivalently as $g(\mathbf{z}(x), y)$ or $g(x, y, \mathbf{z})$.

It is clear that the locations of detected groups within the surveyed strips describe a thinned point process of the original process, namely that of the locations of all the groups within the strips. The detection function operates as the thinning function, deleting groups depending on their location, but as in conventional line transect estimation, it is assumed that deletions occur independently of each other. In practice of course, the validity of this independence assumption will be compromised if the detection of one group affects

the probability of detection of another nearby group. An example of a situation when this might occur is if observers become more alert in the presence of sightings. However, with good survey protocol, such as frequently rotating shifts and appropriate training of observers, the effect of this can be minimised.

An alternative way of viewing the line transect data is as a marked point process, where the detectability of each animal group in the surveyed strip is considered as a characteristic of that group. Thus, each group is marked with some detectability $g_i(x, y)$, $0 \leq g_i(x, y) \leq 1$, $i = 1, \dots, \mathcal{N}$, where \mathcal{N} is the total number of groups available for detection within the surveyed strips. This is an example of position-dependent marking.

The marked point process approach may provide an avenue for the incorporation of sighting-specific covariates, but more methodological development would be required. This is an area that merits further attention, particularly in the light of recent research by Marques (in prep.) on incorporating covariates into the detection function.

2 Model Formulation

2.1 A model based on detection locations

Assuming that animal groups are located according to an inhomogeneous Poisson process of rate $D(x, y)$, the expected number of groups in an area A_i is given by equation (2). Consider two arbitrary points on the trackline $(x_j, 0)$ and $(x_k, 0)$, separated by a distance l along the line. Then the expected number of groups in a rectangular strip between these points and of width w either side of the trackline is

$$\int_{-w}^w \int_{x_j}^{x_k} D(x, y) dx dy. \quad (3)$$

Suppose that some groups located within the strip are not detected. The expected number of detected groups within the strip is a function of how many animals are actually present in the strip and on how detectable they are.

Under the assumption that the detection process is independent of the density of groups,

the locations of detected groups within $\pm w$ of the trackline follow an inhomogeneous Poisson process, with rate $D(x, y)g(x, y)$. This important result — namely that an independent thinning of an inhomogeneous Poisson process is itself an inhomogeneous Poisson process — is proven by the Equivalence Theorem in Cressie (1991), pages 625-6; see also page 690. Given the number of observed detections, n , in the survey area, the assumption that the detections follow an inhomogeneous Poisson process allows us to write down the probability density of the location of an independent random sample from this distribution, as follows:

$$f(x, y) = \frac{D(x, y)g(x, y)}{\int_A D(x, y)g(x, y)dxdy}, \quad (x, y) \text{ in } A, \quad (4)$$

where \int_A denotes that the integral is taken over the survey region (and is equivalent to the double integral $\int_y \int_x$: both notations will be used interchangeably). Given the number of detections, n , the conditional probability density of the locations (x_i, y_i) , $i = 1, \dots, n$, within A is

$$f((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)) = \frac{\prod_{i=1}^n D(x_i, y_i)g(x_i, y_i)}{[\mu(A)]^n}, \quad (5)$$

where $\mu(A) = \int_A D(x, y)g(x, y)dxdy$. Therefore, the joint density of the detection locations and n is

$$f((x_1, y_1), \dots, (x_n, y_n), n) = \begin{cases} \exp[-\mu(A)], & n = 0, \\ \exp[-\mu(A)] \prod_{i=1}^n D(x_i, y_i)g(x_i, y_i)/n!, & n = 1, 2, \dots \end{cases} \quad (6)$$

Although equation (4) follows by definition of an inhomogeneous planar Poisson process, exactly the same result may be derived using an approach analogous to that taken in Seber (1982) in his derivation of the probability density function of perpendicular distances in conventional line transect estimation (see section 3.1).

The expected number of detections in any subarea within the survey area is obtained by integrating the thinned density surface over the corresponding region. Therefore, the expected number of detections, $E(n)$ within a subarea represented by the strip of width $2w$ and length $l = x_k - x_j$ is given by:

$$\begin{aligned}
E(n) &= \int_{-w}^w \int_{x_j}^{x_k} D(x, y)g(x, y)dx dy \\
&= \lambda(x_j, x_k, w) \quad (\text{say}).
\end{aligned} \tag{7}$$

By equation (6), the probability that there are no detections in the strip is

$$P(\text{number of detections in the strip} = 0) = \exp \{-\lambda(x_j, x_k, w)\}. \tag{8}$$

In order that $D(x, y)$ is identifiable from observations of the thinned process $D(x, y)g(x, y)$, a further assumption is required. As in conventional line transect estimation, it is assumed that detection on the trackline is certain, ie. $g(x, 0) = 1, \forall x$. If data are available to estimate $g(x, 0)$, the methodology described in this section can be readily modified to incorporate uncertain detection on the trackline, but we do not consider this further here.

2.2 A likelihood for waiting areas

In the previous chapter, the ‘waiting distance’ was defined as the along-trackline distance between the locations of the sighting platform when detections were made. The random variable M was therefore defined as the length of trackline surveyed before a detection was made. In a sense, this formulation is conditional on the prior estimation of the effective strip width: whilst this is estimated using data on perpendicular distances, the spatial model itself uses data in the along-trackline direction only. This is not an unreasonable approach — the spatial variation in density in the along-trackline direction being much greater than that in the perpendicular distance direction (at least for typical estimated strip widths).

Nevertheless, in this section, a model is formulated which at least in principle, allows spatial variation in density to be modelled both along and perpendicular to the trackline. It is based on the ‘waiting area’ between detections, which would typically be some strip of length equal to the distance between detections, and width to be estimated, but could also be a set of (contiguous) strips of varying width. Within such a framework, covariates affecting detectability which can be continuously recorded along the trackline (such as sea

state) are readily incorporated. In order to simplify the notation, in this derivation it shall be assumed that the width of the strip either side of the transect line is equal, and further, that the width does not vary between detections.

Denoting the location of the start of the surveyed trackline by $\xi_0 = (x_0, 0)$, let \mathcal{A} denote the area surveyed either side of the trackline before a detection is encountered, starting from ξ_0 . The cumulative distribution function (CDF) of \mathcal{A} evaluated at some distance, l , along the trackline from ξ_0 is

$$F_{\mathcal{A}}(l, w|\xi_0) = P(\mathcal{A} \leq 2lw|\xi_0) = 1 - P(\mathcal{A} > 2lw|\xi_0), \quad (9)$$

where w is the distance at which perpendicular distances are truncated, assumed finite. But since $\mathcal{A} > 2lw$ if and only if there were no detections in the strip, then

$$\begin{aligned} F_{\mathcal{A}}(l, w|\xi_0) &= 1 - P(\text{no detections in the strip}) \\ &= 1 - \exp \left\{ - \int_{-w}^w \int_{x_0}^{x_0+l} D(x, y)g(x, y)dx dy \right\} \\ &= 1 - \exp \{ -\lambda(x_0, x_0 + l, w) \}. \end{aligned} \quad (10)$$

Differentiating, we obtain the conditional probability density function of the waiting area, $2lw$, from ξ_0 as follows:

$$\frac{\partial F_{\mathcal{A}}}{\partial l} = \frac{\partial}{\partial l} \left[\int_{-w}^w \int_{x_0}^{x_0+l} D(x, y)g(x, y)dx dy \right] \exp \{ -\lambda(x_0, x_0 + l, w) \}$$

so that

$$f_{\mathcal{A}}(l, w|\xi_0) = \int_{-w}^w D(x_0 + l, y)g(x_0 + l, y)dy \exp[-\lambda(x_0, x_0 + l, w)]. \quad (11)$$

Suppose that there are n detected groups, with actual locations (x_i, y_i) , $i = 1, \dots, n$, and corresponding 'projected locations' $(x_i, 0)$, $i = 1, \dots, n$ (the locations of sightings projected on to the trackline). The first waiting area will be a left-censored observation,

taken as the along-trackline distance from the start of the surveyed trackline to the projected location of the first detection. Similarly, the last waiting area will generally be a right-censored observation — the along-trackline distance from the projected location of the last detection to the end of the surveyed trackline.

Assuming that perpendicular distances are truncated at the same distance w from the trackline throughout the survey area, the conditional joint pdf of waiting areas given ξ_0 may be written:

$$f(l_{n+1}, l_n, l_{n-1}, \dots, l_1, w | \xi_0) = P(\mathcal{A} > 2l_{n+1}w | l_n, \dots, l_1, w, \xi_0) f(l_n, w | l_{n-1}, \dots, w, \xi_0) \dots f(l_1, w | \xi_0) \quad (12)$$

where l_1 is the along-trackline distance from ξ_0 to the projected location of the first detection, $(x_1, 0)$;

l_i , $i = 2, \dots, n$, are the along-trackline distances between the projected locations of the $(i - 1)^{\text{th}}$ and i^{th} detections;

and l_{n+1} is the distance surveyed on effort after the last detection.

Of course, this pdf may be readily simplified to give a joint pdf of the waiting distances, conditional on w , but the pdf was given in the form of equation (12) to emphasise that w is a variable (and its value will clearly have a substantial effect on the observed waiting distances).

Given a vector of spatial parameters $\boldsymbol{\theta}$ for the density surface $D(x, y)$ and parameters $\boldsymbol{\beta}$ for the detection function $g(x, y)$, the conditional likelihood $\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\beta}; \boldsymbol{l}, w | \xi_0)$ is given by:

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\beta}; \boldsymbol{l}, w | \xi_0) = \left[\prod_{i=1}^n \int_{-w}^w D(x_i, y) g(x_i, y) dy \right] \exp \left[- \sum_{i=0}^n \lambda(x_i, x_i + l_{i+1}, w) \right], \quad (13)$$

where \boldsymbol{l} is the vector of observed waiting distances l_1, \dots, l_{n+1} and n is the number of detections. The conditional log-likelihood is therefore:

$$\ln \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\beta}; \boldsymbol{l}, w | \xi_0) = \sum_{i=1}^n \ln \left[\int_{-w}^w D(x_i, y) g(x_i, y) dy \right] - \sum_{i=0}^n \lambda(x_i, x_i + l_{i+1}, w) \quad (14)$$

which, given parametric forms for $D(x, y)$ and $g(x, y)$, can (in principle at least) be maximised numerically.

2.3 A note on maximising the likelihood

In this chapter, we have not attempted to specify a parametric form for either $D(x, y)$ or $g(x, y)$, being more concerned with developing a framework in which spatial modelling from line transect data could take place. In fact, given the data available, maximisation of the likelihood in equation (13) is unlikely to be straightforward, unless (as in the next section), the model is simplified substantially. The data that one would typically obtain from a line transect survey in order to maximise this likelihood would be distances between detection locations, and the perpendicular sighting distance of each detection, together with auxiliary spatial covariates \mathbf{z} , such as latitude and longitude, and auxiliary ‘sightability’ covariates such as Beaufort sea state. The two sets of covariates are not necessarily mutually exclusive. For example, in the Southern Ocean, the sea state close to the ice edge tends to be lower, therefore detectability of whales would be expected to decrease with increasing distance from the ice edge (or say, increasing latitude for an ice edge configured east-westerly), all other things being equal. Research is currently being undertaken to examine methods which can be used to implement the approach proposed in this chapter, but this will not be completed in time for inclusion in this thesis. Two methods of implementation are under investigation: the first uses the EM-algorithm (Dempster *et al.*, 1977) and the second uses a Markov Chain Monte Carlo simulation approach (see, for example, Gamerman (1997)).

2.4 Maximisation of the likelihood: a simplified case

Consider the following simplified case:

- density varying only along the transect lines, ie. $D(x, y) = D(x)$, independently of y , for $|y| \leq w$
- a univariate detection function, $g(y)$, depending only on perpendicular distance, y , from the transect line.

Under these restrictions, the likelihood function simplifies to

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\beta}; \mathbf{l}, w | \xi_0) = \left[2\mu \prod_{i=1}^n D(x_i) \right] \exp \left[-2\mu \sum_{i=0}^n \int_{x_i}^{x_i+l_{i+1}} D(x) dx \right], \quad (15)$$

where $\mu = \int_0^w g(y) dy$ is the effective strip half-width. Given parametric forms for $D(x)$ and $g(y)$, the parameter estimates which maximise this likelihood can be obtained using standard numerical maximisation routines. The likelihood above was maximised for the minke whale data from the IWC/IDCR 1992-3 Antarctic survey (as used in the example applications in chapters 3 and 4) using the ‘amoeba’ routine of Press *et al.* (1986). This routine implements the downhill simplex method of Nelder and Mead (1965).

In practice, the assumption that $D(x, y)$ is independent of y for $|y| \leq w$ will be reasonable for most applications, because typically detection distances are small relative to the lengths of the lines, and relative to the scale at which animal density varies. The assumption that $g(x, y)$ is independent of x is a common assumption in line transect estimation, relying on the property of ‘pooling robustness’ (Burnham *et al.*, 1980) for modelling $g(y)$.

3 Some comparisons with conventional line transect estimation

3.1 A pdf for detection locations

In the previous section, a pdf for detection locations (x_i, y_i) , $i = 1, \dots, n$ was obtained by assuming that the locations followed an inhomogeneous Poisson process with spatially varying rate, $D(x, y)g(x, y)$. In this section, this pdf is derived from a line transect perspective, following Seber (1982).

If the density of animal groups at location (x, y) is given by $D(x, y)$, then the expected number of groups ‘at’ location (x, y) is given by the number of groups in some small area da , defined by co-ordinates at (x, y) , $(x + dx, y)$, $(x, y + dy)$ and $(x + dx, y + dy)$ as $D(x, y) dx dy$. The number of groups throughout the survey area, A , is the integral of all such areas in the survey region, $\int_A D(x, y) dx dy$. Then if a group is located somewhere in

A , the probability that it is located in da is

$$P[\text{the group is located in } da] = \frac{D(x, y) dx dy}{\int_A D(x, y) dx dy}.$$

Thus, whereas the conventional line transect approach would assume that the probability of a group being located at (x, y) is simply the probability that that location is included in the sample (i.e. $dx dy/A$), here the corresponding probability is also dependent upon the density of groups at that location. Since the conditional probability of a group being detected given that it is at location (x, y) is by definition, $g(x, y)$, it follows that

$$P[\text{the group is detected in } da] = \frac{D(x, y) g(x, y) dx dy}{\int_A D(x, y) dx dy}.$$

The probability that a group is seen is

$$P[\text{a group is detected}] = \frac{\int_y \int_x D(x, y) g(x, y) dx dy}{\int_A D(x, y) dx dy}.$$

Since

$$\begin{aligned} P[\text{the group is located in } da \mid \text{a group is detected}] &= \frac{P[\text{the group is detected in } da]}{P[\text{a group is detected}]} \\ &= \frac{D(x, y) g(x, y) dx dy}{\int_y \int_x D(x, y) g(x, y) dx dy}, \end{aligned}$$

then

$$f(x, y) dx dy = \frac{D(x, y) g(x, y) dx dy}{\int_y \int_x D(x, y) g(x, y) dx dy},$$

and the pdf given in equation (4) follows immediately. If $D(x, y) = D \forall (x, y) \in A$, the pdf of detection locations is given simply as the spatial equivalent of the familiar line transect pdf of perpendicular distances, i.e. $f(x, y) = g(x, y) / \int_y \int_x g(x, y) dx dy$, cf. equation (2.7) of Seber (1982).

3.2 Equivalence of the point process likelihood derivation and the conventional line transect estimator

In section 2.4, the likelihood given in equation (13) was simplified to enable it to be fitted using standard maximisation routines. This allowed estimation of a univariate detection

function, $g(y)$ (assumed symmetric), and estimation of a density function $D(x)$ which depended on spatial covariates (latitude, longitude and distance from the ice edge). Conventional line transect analyses estimate only a single average density for each stratum so that the $D(x, y)$ of this chapter, with corresponding parameter vector θ , is equivalently represented by a single parameter to be estimated, D .

The log-likelihood (equation 14) thus simplifies to

$$\ln\mathcal{L}(D, \beta; \mathbf{l}, w|\xi_0) = n \ln \left[2D \int_0^w g(y) dy \right] - 2DL \int_0^w g(y) dy, \quad (16)$$

where L is the total distance surveyed. Differentiation of equation (16) with respect to D yields the maximum likelihood estimate of density, \hat{D}_{ML} (say). As shown below this is equivalent to the conventional line transect estimate of group density:

$$\frac{\partial \ln \mathcal{L}}{\partial D} = \frac{n}{D} - 2L \int_0^w g(y) dy$$

so that

$$\hat{D}_{ML} = \frac{n}{2L \int_0^w \hat{g}(y) dy}. \quad (17)$$

4 Discussion and further work

The methodology proposed in this chapter is incomplete. It does, however, establish a theoretical grounding for spatial modelling from line transect data, and perhaps most importantly, the methodology has been developed from an applied perspective. Previous accounts noting the possibility of formulating the line transect method as a point process (Stoyan, 1982; Högmander, 1991; Högmander, 1995) were not directed towards actually using the point process approach to model line transect data in a practical sense. In particular, attention was restricted to stationary processes — a somewhat unrealistic situation in most ecological applications. Nevertheless, their work represents a significant contribution to the literature linking point process theory and line transect estimation.

A number of possible areas for future development of the modelling framework presented in this chapter have arisen during the course of this work. The formulation itself is a first

step in attempting to account fully for spatial variation in both the detection function and the density of animal groups in estimation of abundance. In future work, we plan to implement the full approach, using for example the EM algorithm or MCMC simulation approaches mentioned earlier. In the interim, the modelling proposed in chapter 4, utilizing standard statistical software, provides a practical means to fit spatial models to waiting distance-type data, although such modelling is necessarily conditional on the estimated detection function.

Chapter 6

Modelling heterogeneity in mark-recapture analyses

1 Introduction

In the present thesis, attention has focussed largely on line transects, and in particular, the primary aim has been to develop a spatial framework for modelling line transect data. In this chapter, attention is turned to another widely used method for the assessment of wildlife populations: mark-recapture analyses. However, there is no intention in this chapter to review either the vast literature or the wide array of models that are available for use in mark-recapture studies; references will only be made when they are of direct relevance to the research presented here.

The work arose from a specific case study, carried out in collaboration with Dr Ben Wilson, involving the estimation of survival rates and abundance of a coastal population of bottlenose dolphins (*Tursiops truncatus*) in the Moray Firth, Scotland.

The methodology presented in this chapter is based on earlier research carried out independently by Huggins (1989; 1991) and Alho (1990), together with an application that used their approach (hereafter termed the ‘Alho-Huggins’ approach) by Buckland *et al.* (1993b). The Alho-Huggins methodology was originally developed in order to estimate the size of a closed population when there is heterogeneity in the capture probabilities of individuals. Both Huggins (1989; 1991) and Alho (1990) implemented this methodol-

ogy by direct maximisation of a conditional likelihood (based on the captured individuals only), using general purpose optimization routines. A brief description of their approach is given in section 2.1.

Buckland *et al.* (1993b) adopted the Alho-Huggins approach to analyse data from grey whale migration counts, in which animals were recorded from two look-out stations as they passed by cliffs off Monterey, California. These data may be viewed as though they were from a mark-recapture survey with two sampling occasions. Buckland *et al.* (1993b) showed how the Alho-Huggins methodology could be implemented with standard statistical software, using logistic regression to model the conditional capture probabilities, and including an offset in the linear predictor. However, their approach is quite specific to mark-recapture data with only two capture occasions and does not use data on the past capture histories of individuals. It can be shown that inclusion of the past capture histories in the conditional probabilities requires exclusion of the data from all individuals seen only on the last sampling occasion, which for the case of two sampling occasions, is clearly not a desirable option. Algebraic details are given in Appendix A.

In section 2.2, it is shown that the methodology presented in Buckland *et al.* (1993b) may be extended to allow models to be fitted to data from more than two sampling occasions, using standard statistical software. Past capture histories are incorporated into the modelling framework, although as noted above, this is only reasonable when the number of sampling occasions exceeds two.

In section 3.1, the methodology is applied to the Moray Firth bottlenose dolphin data, for which it is thought reasonable to assume the population is closed throughout the summer months (from May to September), and thus the number of sampling occasions in this example is five (although with sufficient persistence with the algebra, the methodology can be applied to t sampling occasions, where $t > 2$). The estimates resulting from this modelling are compared with estimates from Wilson *et al.* (1999). However, because of discrepancies between the data used here and those used by Wilson *et al.*, the estimates are not directly comparable.

2 Modelling heterogeneity in capture probabilities

2.1 The Alho-Huggins approach

The Alho-Huggins approach (Huggins, 1989, 1991; Alho, 1990) provides a method for the incorporation of observable covariates in a closed population mark-recapture analysis. It is described briefly here.

Let p_{ij} denote the probability that individual i is captured on sampling occasion j , where $i = 1, \dots, N$ are the individuals in the population (N denoting the size of the population), and $j = 1, \dots, t$ are the sampling occasions. Let the past capture history of individual i up to sampling occasion j be denoted by z_{ij} , where

$$z_{ij} = \begin{cases} 1 & \text{if individual } i \text{ has been captured before occasion } j \\ 0 & \text{otherwise} \end{cases} .$$

Thus, π_{ij} , the conditional probability that individual i is captured on sampling occasion j given its past capture history and that it is captured at least once, is given by

$$\pi_{ij} = \frac{p_{ij}}{1 - (1 - z_{ij}) \prod_{k=j}^t (1 - p_{ik})} . \quad (1)$$

Given a parameter vector, β , and covariates representing the characteristics of the individuals and/or differences in the environmental conditions on each sampling occasion, then the logistic model can be used to model the p_{ij} as a function of the covariates. Inferences may then be based on the following conditional likelihood, which itself is based only on the n captured individuals:

$$\mathcal{L}(\beta; \mathbf{y}) = \prod_{i=1}^n \prod_{j=1}^t \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{(1-y_{ij})} , \quad (2)$$

where

$$y_{ij} = \begin{cases} 1 & \text{individual } i \text{ is captured on occasion } j \\ 0 & \text{otherwise} \end{cases} .$$

Due to the non-standard form of the π_{ij} appearing in the conditional likelihood shown in equation (2) above, both Huggins (1989; 1991) and Alho (1990) utilized general purpose optimization routines for maximising the likelihood, using logistic regression to model the p_{ij} .

As noted in section 1, the Alho-Huggins approach was adopted by Buckland *et al.* (1993b) in order to model heterogeneity in the detection probabilities of migrating grey whales, from a shore-based count. They used a simplified form of equation (1), in which the probabilities, π_{ij} , are not conditional on the past capture histories, z_{ij} , of each detected individual, but are conditional on each detected individual i , $i = 1, \dots, n$, being detected. Noting that the logit of these conditional probabilities could be expressed in terms of the p_{ij} as

$$\ln \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \ln \left[\frac{p_{ij}}{(1 - p_{ij})} \right] - \ln(p_{ij'}), \quad j = 1, 2 \text{ and } j' = 3 - j, \quad (3)$$

Buckland *et al.* (1993b) used logistic regression to model the π_{ij} directly. Consequently, the term $-\ln(p_{ij'})$ was included as an offset in the linear predictor.

In the next section, an extension of the Buckland *et al.* methodology is proposed, which incorporates past capture history data when there are more than two sampling occasions, but unlike the implementations presented in Huggins (1989; 1991) and Alho (1990), may still be fitted using standard statistical software.

2.2 A general logistic regression approach for modelling capture probabilities

In Appendix A, it is shown algebraically that the inclusion of past capture history data when there are only two sampling occasions (as in the grey whale analysis of Buckland *et al.*, 1993b) results in an ill-defined form for p_{ij} when an individual is not seen on the first occasion ($y_{i1} = 0$, $z_{i1} = 0$) but is seen on the second occasion ($y_{i2} = 1$, $z_{i2} = 0$).

However, when there are more than two sampling occasions, past capture history data may be included, and an approach using logistic regression of the conditional capture probabilities can be used. An offset term must be included to adjust for modelling the

conditional, rather than unconditional, probabilities.

It is assumed that the probability of capturing an individual on any given sampling occasion is independent of whether it is captured on any other occasion. It is also assumed that the capture probability of an individual on a given sampling occasion is independent of the capture probability of other individuals on that occasion.

Suppose that the number of sampling occasions is given by t , $t > 2$. Given the past capture histories of the captured individuals (and implicitly that these individuals have been captured at least once), then the probability, π_{ij} , that the i^{th} individual is captured on the j^{th} sampling occasion was given in equation (1) as

$$\pi_{ij} = \frac{p_{ij}}{1 - (1 - z_{ij}) \prod_{k=j}^t (1 - p_{ik})}.$$

In order to model the π_{ij} , $i = 1, \dots, n$, $j = 1, \dots, t$, directly within a logistic regression framework, the logit of these conditional probabilities must be expressed in terms of the logit of the unconditional probabilities (the p_{ij}) plus, if necessary, some other term (which becomes an offset in the linear predictor). Unlike the Buckland *et al.* (1993b) application, by incorporating the past capture histories of each individual in the modelling, the form of the offset term is different for each sampling occasion. In addition, the algebra is more complicated, and becomes increasingly so as the number of sampling occasions increases. However, the offset is always tractable, and a symbolic computation package, such as Maple (Waterloo Maple Inc., 1994), may be used to obtain an expression for it, should the algebra become too tedious.

To avoid filling the next few pages with algebra, we consider the case $t = 3$. Noting that $z_{i1} = 0 \ \forall i$, the conditional probabilities of capture of individual i on the first, second and third sampling occasions are given respectively by:

$$\pi_{i1} = \frac{p_{i1}}{1 - (1 - p_{i1})(1 - p_{i2})(1 - p_{i3})};$$

$$\pi_{i2} = \frac{p_{i2}}{1 - (1 - z_{i2})(1 - p_{i2})(1 - p_{i3})}; \quad \text{and}$$

$$\pi_{i3} = \frac{p_{i3}}{1 - (1 - z_{i3})(1 - p_{i3})}.$$

From these equations, the following expressions for the unconditional probabilities of capture may be obtained:

$$p_{i1} = \frac{\pi_{i1}(z_{i2}\pi_{i2}\pi_{i3} - z_{i2}\pi_{i2} - z_{i3}\pi_{i3})}{\pi_{i1}\pi_{i2}\pi_{i3} - \pi_{i1}\pi_{i2} - \pi_{i1}\pi_{i3} - \pi_{i2}\pi_{i3} + \pi_{i1} + \pi_{i2} + \pi_{i3} - z_{i2}\pi_{i2} - z_{i3}\pi_{i3} + z_{i2}\pi_{i2}\pi_{i3} - 1};$$

$$p_{i2} = \frac{\pi_{i2}(z_{i2}\pi_{i3} - z_{i3}\pi_{i3} - z_{i2})}{z_{i2}\pi_{i2}\pi_{i3} - \pi_{i2}\pi_{i3} + \pi_{i2} + \pi_{i3} - z_{i2}\pi_{i2} - z_{i3}\pi_{i3} - 1};$$

$$\text{and } p_{i3} = \frac{z_{i3}\pi_{i3}}{1 - \pi_{i3} + z_{i3}\pi_{i3}}.$$

Now let $f_{ij} = \ln[\pi_{ij}(1 - \pi_{ij})^{-1}]$ and $g_{ij} = \ln[p_{ij}(1 - p_{ij})^{-1}]$, then the following equations express the g_{ij} in terms of the z_{ij} and π_{ij} :

$$\begin{aligned} g_{i1} &= \ln \left[\frac{\pi_{i1}(z_{i2}\pi_{i2}\pi_{i3} - z_{i2}\pi_{i2} - z_{i3}\pi_{i3})}{(1 - \pi_{i1})(\pi_{i1}\pi_{i2}\pi_{i3} - \pi_{i1}\pi_{i2} - \pi_{i1}\pi_{i3} - \pi_{i2}\pi_{i3} + \pi_{i1} + \pi_{i2} + \pi_{i3} - z_{i2}\pi_{i2} - z_{i3}\pi_{i3} + z_{i2}\pi_{i2}\pi_{i3} - 1)} \right] \\ &= f_{i1} + \ln(\text{off}_{i1}); \end{aligned}$$

$$\begin{aligned} g_{i2} &= \ln \left[\frac{\pi_{i2}(z_{i2}\pi_{i3} - z_{i3}\pi_{i3} - z_{i2})}{(1 - \pi_{i2})(\pi_{i3} - z_{i3}\pi_{i3} - 1)} \right] \\ &= f_{i2} + \ln(\text{off}_{i2}); \end{aligned}$$

$$\begin{aligned} \text{and } g_{i3} &= \ln \left[\frac{z_{i3}\pi_{i3}}{(1 - \pi_{i3})} \right] \\ &= f_{i3}, \end{aligned}$$

where off_{ij} , $j = 1, 2$, are the offset variables on the first and second sampling occasions respectively. Note that the offset for the third sampling occasion is zero, since we are assuming that $z_{i3} = 1 \forall i$. Data for which any of the $z_{i3} = 0$ (ie. any individuals seen only on the last sampling occasion) must be excluded from the analysis. This option was ruled out for the two-sample mark-recapture case, and it is not ideal when there are only three capture occasions either, but it becomes less of a problem as the number of sampling occasions increases.

The expressions given above may look frightening, but they are numerically simple to evaluate, and in practice, depending on whether the values of z_{i2} and z_{i3} are 1 or 0, some of the terms will disappear.

In the next section, an algorithm for implementing the methodology in order to obtain closed population estimates of abundance using any standard logistic regression package is given.

2.3 An algorithm for fitting the logistic regression model

The mark-recapture data required to implement the methodology described in the previous section comprise:

- observations y_{ij} , $i = 1, \dots, n$ (the number of captured individuals), and $j = 1, \dots, t$ (the number of sampling occasions), where $y_{ij} = 1$ or 0 , indicating whether the captured individual i is seen, or not seen, on occasion j ;
- past capture histories z_{ij} , $i = 1, \dots, n$, $j = 1, \dots, t$, where $z_{i1} = 0 \ \forall i$ (these values are calculated from the y_{ij});
- observable covariates which may be related to the individuals or to the environmental conditions observed on each sampling occasion — in either case, they can be categorical or continuous covariates.

The following procedure may be used to fit a logistic regression model to the y_{ij} , and thus obtain an estimate of abundance which accommodates heterogeneity in capture probabilities. It is assumed that data from any individuals seen only on the last capture occasion are excluded from the analysis.

1. Fit a logistic regression model to the observed y_{ij} , with no offset term.
2. Using the fitted values from step 1 (call these $\tilde{\pi}_{ij}^{(0)}$), estimate an offset value for each individual on each sampling occasion, j , $j = 1, \dots, t$. (For $t = 3$, expressions for determining the offset are given in section 2.2; replace π_{ij} by $\tilde{\pi}_{ij}^{(0)}$ in the expressions for the g_{ij} .)
3. Refit the logistic model, including the estimated offset in the linear predictor, to obtain new fitted values, $\tilde{\pi}_{ij}^{(1)}$.
4. Using the new values of $\tilde{\pi}_{ij}^{(1)}$, repeat steps 2 and 3 until convergence is reached.

5. Having achieved convergence of the offset values, refit the logistic model with the offset, but now allow stepwise covariate selection.
6. Using the fitted values of the model selected in step 5, estimate the probability, p_i , that each individual was seen at least once. For individual i , this is given by

$$\hat{p}_i = 1 - \prod_{j=1}^t (1 - \hat{p}_{ij})$$

and the \hat{p}_{ij} are evaluated from the final fitted values $\tilde{\pi}_{ij}$. (For $t = 3$, expressions for the p_{ij} are given in section 2.2; replace π_{ij} by $\tilde{\pi}_{ij}$.)

7. Estimated abundance, \hat{N} of the population is given by

$$\hat{N} = \sum_{i=1}^n \frac{1}{\hat{p}_i}$$

(Huggins (1989; 1991), Alho (1990)).

An S-PLUS program for implementing this algorithm is given in Appendix B. Note that following covariate selection (step 5), it may in some circumstances (for example, if the model is substantially different from the one used to estimate the offset), be preferable to re-estimate the offset. This is straightforward — simply iterate steps 3, 4 and 5 until the covariates selected do not change (and the offset has converged).

3 Application of the methodology

3.1 A photo-identification study of bottlenose dolphins

3.1.1 Background

It is widely believed that the number of bottlenose dolphins in British waters has declined in recent decades (Kayes, 1985). Particular concern has been expressed over the status of a resident population in the Moray Firth, Scotland, which inhabits waters at the northern extent of the range of this species, and is the only known population remaining in the North Sea (Wilson *et al.*, 1999). A high prevalence of skin lesions (Wilson *et al.*, 1997b), use of waterways polluted by human sewage (Curran *et al.*, 1996), and vessel traffic, including

a growing number of dolphin-watching boats (Janik and Thompson, 1996), have led to a demand for information on the possible impacts on the population.

In order to implement an effective conservation strategy, an estimate of the population size and status was required. The bottlenose dolphins inhabiting the Moray Firth possess a number of natural markings which can be used to identify individuals, ranging from permanent features, such as dorsal fin nicks and humped backs, to temporary features such as minor body wounds (Wilson, 1995). Therefore, in the summer of 1989, a photo-identification study of the dolphins inhabiting the Moray Firth was initiated. Results from data collected from surveys between 1990 and 1992 were presented in Wilson *et al.* (1999). Using a closed population model which allowed capture probabilities to vary by time and by individual (Model M_{th} of the program CAPTURE (Otis *et al.*, 1978)), the number of individuals (excluding calves) for 1992 was estimated to be 129 (C.V. = 0.12). Changes in size of a population of this number would be expected to be slow (Wilson *et al.*, 1999). Coupled with large confidence intervals around the estimate, this results in low power to detect future trends in population size. Open population models were not considered by Wilson *et al.* (1999) since they cannot take account of the heterogeneity in individual sighting probabilities observed and modelled under the assumption of closure.

3.1.2 The study area and a description of the data

Surveys were conducted in the inner and outer regions of the Moray Firth (Figure 1), but logistically it was only possible to conduct surveys in the outer area infrequently. The dolphins were observed to exhibit a seasonal distribution in the Firth (Wilson *et al.*, 1997a), and all individuals seen in the outer area during October to April were also seen in the inner area during May to September (Wilson, 1995). Therefore, a reliable estimate of population size may be calculated from the data from the May to September surveys in the inner area only. During these months, it is thought reasonable to assume that the population is demographically closed.

The sampling protocol on encountering a school of dolphins was to photograph as many animals as possible in the school, from both the left and right sides. Effort was taken to ensure that neither the most gregarious individuals, nor those with the most obvious markings, were more likely to be photographed. The matching process included catego-

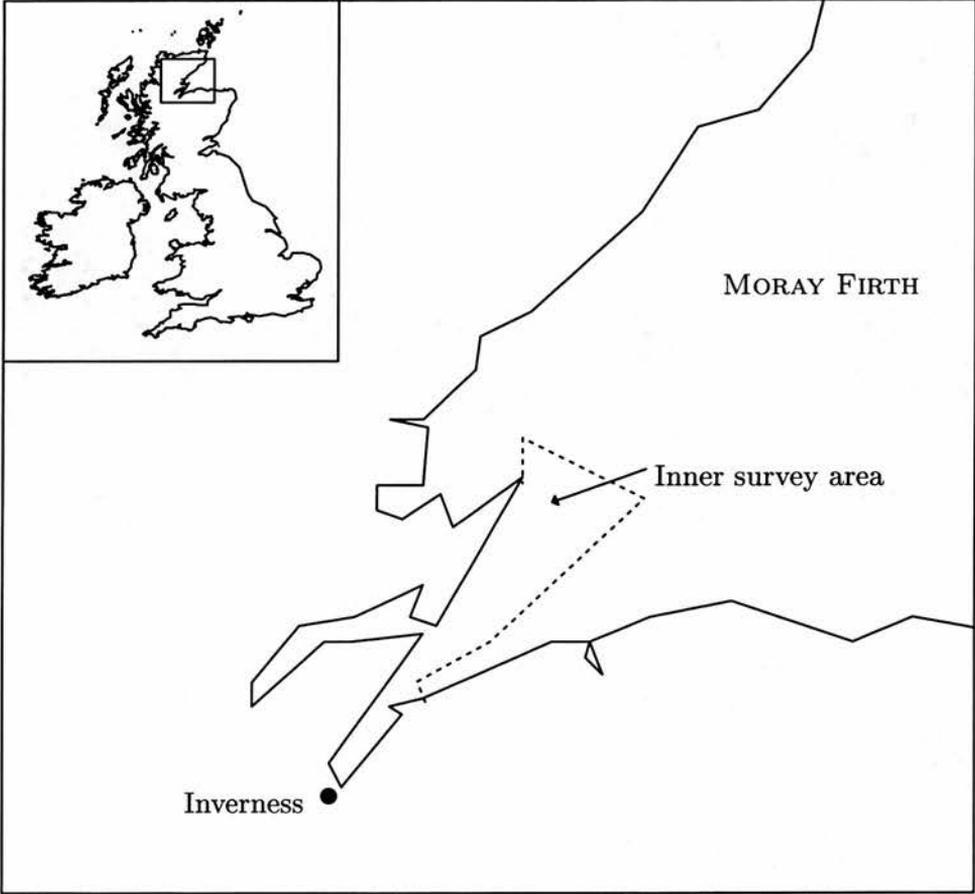


Figure 1: Area within the inner region of the Moray Firth where the surveys took place.

rization of the photographs by quality, and classification of the identifiable features into those which are long-lasting and those which are only temporary. ‘Grade 3’ photographs were well-lit, in focus, free from spray, and taken parallel with the exposed flank of the individual such that if any patches of skin with active disease were present (the most subtle marking used for identification), they would be visible. All dolphins encountered in the Moray Firth (except some very young animals) possess sufficient markings that they could be identified from a grade 3 photograph, and it is considered that photographs of a lower quality should not be included in an analysis (Wilson *et al.*, 1999). Data from grade 3 photographs were only available for the years 1990-1992.

Following Wilson *et al.* (1999), the data used in this analysis were obtained from grade 3 photographs of individuals with long-lasting marks. The data are separated — and analysed separately — according to whether the photographs were from the left or the right sides of the animals. Assuming closure of the population between May and September, there are therefore five sampling occasions — each month being considered as a single sample (although there would generally have been more than one survey per month). During these five summer months, it is assumed that the population is closed. Furthermore, the probability of photographing a given individual in a given summer month is assumed to be independent of whether it is photographed in any other summer month, and independent of the probability of photographing any other individual that month.

Past capture histories were calculated from the sightings data. The age category (adult, reproductive female, sub-adult, juvenile, calf or neonate) of the individuals was the only available covariate from the data collected, but the factor ‘month’ and the past capture histories were also considered for inclusion in the logistic regression model. Since the probability of photographing a calf or neonate is not independent of the probability of photographing its mother (Wells and Scott, 1990), data from these animals were excluded from the analysis.

3.1.3 Results from fitting the general logistic model

The methodology described in section 2.2 was applied to the Moray Firth bottlenose dolphin data. Model fitting was carried out in S-PLUS. Initial fitted values, $\tilde{\pi}_{ij}$, of the conditional probabilities, from which the offset was estimated, were obtained by fitting a ‘full’

logistic model (ie. one with all potential covariates included) to the monthly sightings data from the 1990–92 surveys, separately for left and right photographs. Potential explanatory variables were age and month (fitted as factors), and past capture history. (Maple code which provides the expressions required to estimate the offset and the unconditional capture probabilities is given in Appendix C.) In estimating the offset, convergence was rapid, usually being achieved within five iterations. The model was refitted once convergence was achieved allowing selection of explanatory variables. In all cases, the fitted model was

$$E \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \exp \{ \alpha + \beta_j + \gamma z_{ij} \}, \quad i = 1, \dots, n, \quad j = 1, \dots, 5, \quad (4)$$

where α is the intercept parameter, β_j are parameters representing the factor month, and γ is the coefficient of the covariate z_{ij} .

Year	Side	$\hat{\theta}$	Model M_{th}				Logistic Regression Model			
			\hat{N}_{long}	%CV	\hat{N}_{tot}	%CV	\hat{N}_{long}	%CV	\hat{N}_{tot}	%CV
1990	L	0.68	100	26.2	147	27.0	124	11.2	182	13.0
1990	R	0.65	76	16.1	117	17.1	92	6.1	141	8.6
1991	L	0.58	45	8.2	78	10.3	54	3.9	93	7.4
1991	R	0.56	49	8.0	88	10.0	58	3.7	104	7.0
1992	L	0.57	84	23.9	147	24.2	68	8.1	120	11.3
1992	R	0.60	59	15.1	98	16.5	59	4.1	99	7.9

Table 1: Comparison of estimates of abundance of the inner Moray Firth bottlenose dolphin population (a) taken from Wilson *et al.* (1999) using Model M_{th} of program CAPTURE (Otis *et al.*, 1978) and (b) using the logistic regression approach described in this chapter. Data are from left (L) and right (R) photographs from 1990–92. $\hat{\theta}$ is the estimated proportion of dolphins with long-lasting marks in the population; \hat{N}_{long} is the estimated number of dolphins with long lasting marks; and \hat{N}_{tot} is the estimated abundance of the population.

For each year, the estimated abundance (\hat{N}_{long}) from fitting the logistic model separately to the left and right photographs is given in Table 1. (The ‘long’ subscript indicates that the estimates were calculated using data from animals with long-lasting marks only.) The corresponding CVs were calculated using a simplified version of the variance estimator as

originally given in Huggins (1989; 1991) and Alho (1990). Their estimator incorporates a component of the variance due to estimation of the parameters of the model; the following expression for the variance of \hat{N}_{long} assumes that these parameters are known:

$$\text{var}(\hat{N}_{long}) = \sum_{i=1}^n \frac{(1 - \hat{p}_i)}{\hat{p}_i^2}.$$

This estimator was used by Buckland *et al.* (1993b) in their analysis of grey whale data. Buckland (pers. comm.) notes that when the probability of capture (p_i) is close to unity, the component of the variance due to estimation of the parameters is relatively small. In our example, the average value of \hat{p}_i was quite high, particularly in 1991 and 1992, exceeding 0.9 for three of the four cases. Nevertheless, it should be noted that the CVs presented in Table 1 from the logistic regression model are negatively biased, because the uncertainty due to estimating the model parameters has not been incorporated. However, Huggins (1989) comments that this source of uncertainty is not incorporated in the variance estimator used in the CAPTURE models of Otis *et al.* (1978) either, so at least for comparison purposes, the results presented in Table 1 in which the logistic regression estimates are compared with those from Model M_{th} in a CAPTURE analysis (Wilson *et al.*, 1999) are informative. Since the data were from animals with long-lasting marks only, an estimate of the total number of animals in the population (\hat{N}_{tot}) may be calculated using the estimated proportion of animals in the population with such marks (Wilson *et al.*, 1999). This assumes that the capture probabilities are independent of whether or not the individuals had long-lasting marks. These results are also given in Table 1.

4 Discussion

Whilst it is not possible to draw any firm conclusions from applying the methodology to just one data set, the results from implementation of the approach with the Moray Firth bottlenose data indicate that the method yields estimates which are comparable with those from conventional closed population models. Direct comparisons with the estimates from Wilson *et al.* (1999) are not valid, since the data used in their analysis are more recent than those used in the analysis presented here. However, as a rough gauge to the performance of the method, one may examine the within-year comparisons of the estimates in Table 1 between left and right-sided photographs. The differences between the left and right-sided

estimates from the logistic regression model were 41, 11 and 21 animals for 1990, 1991 and 1992 respectively, or as a percentage of the mean of the two estimates for each year, 25%, 11% and 19%. Because the estimates from the logistic regression model are generally higher than those from the CAPTURE analysis of Wilson *et al.* (1999), the percentage differences provide a better comparison of the differences between the estimates from left and right-sided photographs between the two approaches. The corresponding percentage differences from the Model M_{th} estimates are 23%, 12% and 20%, indicating a broadly similar performance between the two approaches in this regard.

Inspection of the logistic regression estimates in Table 1 reveals substantial differences between the estimates of abundance for 1990 and those for other years, particularly for the left-sided photographs. It is not possible to conclude whether the relatively high estimated abundances in 1990 from the logistic regression method, particularly for the left-sided data, are due to a failing of the method itself; they may be caused by a real effect, or they may arise simply as an artefact of the data. The suggestion that the effect might be real is not infeasible, for example, there may have been a temporary influx of animals from the outer Moray Firth in that year. (The average abundance from the 1990 Model M_{th} estimates was also higher than that in other years, providing some evidence to support this possibility.) The data from 1990 are thought to be reliable, the survey protocols having been largely established the previous year. However a high proportion of the dolphins encountered in 1990 were seen only once in the summer months, producing anomalous capture history data compared to the subsequent years. It is not uncharacteristic for data in such long-term studies to show irregularities at the beginning of a time series, and whilst the survey protocol was more established than in the first year of the study, it is possible that it was not completely standardized until the following year. Furthermore, some temporal changes would be expected as familiarity with the established procedures increased.

The increased precision of the estimates gained by the logistic regression model compared with the precision of the estimates from the CAPTURE analysis is encouraging. Although the coefficients of variation (CVs) from the logistic regression model are similar to those from the CAPTURE analysis in 1991, notably they are approximately half the value of those from CAPTURE in 1990 and 1992. A gain in precision of this amount would be expected to require about a fourfold increase in effort. Although further investigation into the performance of the logistic regression approach is still required, in particular to

examine whether the effect of incorporating the component of variance due to estimation of parameters differs between methods, the gain in precision seen in this example indicates a potentially useful approach for increasing the power to detect a trend in population estimates across years.

In fact, one of the aims of this research at the outset of the collaboration was to examine the potential use of open population models in order to examine the dynamics of the population in more detail than would be possible from serial closed population estimates. However, conventional open population models (Jolly, 1965; Seber, 1965) are unable to incorporate heterogeneity in individual sighting probabilities observed and modelled by Wilson *et al.* (1999) under the assumption of closure. The ‘robust design’ (Pollock, 1982) has recently received much attention (see, for example, Kendall and Pollock (1992); Kendall *et al.* (1995) and references therein). The design consists of a series of ‘primary’ sampling periods which are widely spaced in time (years, for example), within which there are a series of ‘secondary’ sampling periods sufficiently closely spaced in time so that it is reasonable to assume that the population under study is closed (Figure 2). It allows the use of both closed and open population models for long-term mark-recapture studies.

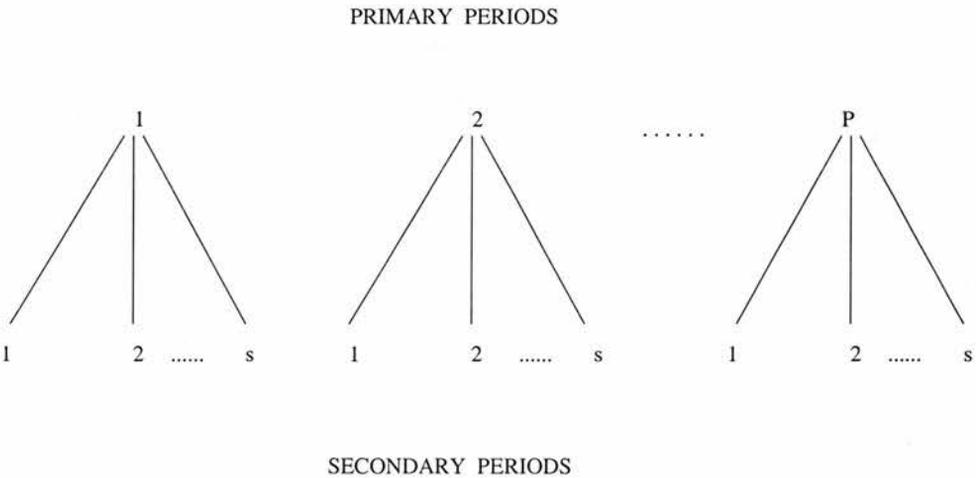


Figure 2: The robust design of Pollock (1982), shown applied to P primary sampling periods and within each of these, s secondary sampling periods.

An obvious extension to the analysis presented in section 3.1 is to combine the robust design with the logistic regression methodology to provide estimates from a type of robust open population modelling which also incorporate heterogeneity in sighting probabilities. For the Moray Firth bottlenose dolphin data, the primary sampling occasions would be

years and appropriate secondary sampling periods would be months.

In conclusion, the methodology presented in this chapter provides a general means for estimating closed population estimates of abundance from mark-recapture data when there is heterogeneity in capture probabilities. Past capture histories may be incorporated; this may be particularly useful when the individuals are thought to exhibit 'trap-happy' or 'trap-shy' responses, although this is obviously not generally considered to be the case for photo-identification studies. The method is easy to implement, and the models may be fitted using standard statistical software. In this section, the model was fitted within a generalized linear modelling framework which was adequate for the Moray Firth bottlenose data since there were few explanatory variables, and these were categorical. The methodology is sufficiently general that it may be extended to a generalized additive modelling framework; in practice, this is of most use when at least one continuous covariate is available, and further, that the relationship between the logit of the capture probabilities and such a covariate is nonlinear.

Appendices

Appendix A: Inclusion of past capture histories when there are two sampling occasions

The conditional capture probabilities, π_{ij} , used to model grey whale migration counts by Buckland *et al.* (1993b) do not use past capture histories, z_{ij} , as proposed by Huggins (1989; 1991). It is shown below that including the past capture histories when there are only two sampling occasions results in the unconditional probability of capture on the second sampling occasion, p_{i2} , being unestimable from the conditional probabilities when $z_{i2} = 0$, ie. when an individual is not seen on occasion 1 and is seen on occasion 2.

As in Huggins (1989; 1991), define the probability of capture of the i^{th} individual on the j^{th} sampling occasion, given the past capture histories and that the individual is seen at least once, to be

$$\pi_{ij} = \frac{p_{ij}}{1 - (1 - z_{ij}) \prod_{k=j}^t (1 - p_{ik})}.$$

Now since $z_{i1} = 0 \forall i$, the capture probabilities on the first and second sampling occasions are given by:

$$p_{i1} = \frac{z_{i2}\pi_{i1}\pi_{i2}}{1 - \pi_{i1} - \pi_{i2} + \pi_{i1}\pi_{i2} + z_{i2}\pi_{i2}} \quad (5)$$

and

$$p_{i2} = \frac{z_{i2}\pi_{i2}}{1 - \pi_{i2} + z_{i2}\pi_{i2}}. \quad (6)$$

Thus $p_{i1} = 0$ if $z_{i2} = 0$ (by definition this is true), and if $z_{i2} = 1$, p_{i1} can be expressed in terms of the π_{ij} , $j = 1, 2$. If $z_{i2} = 1$, p_{i2} may be defined in terms of π_{i2} . If $z_{i2} = 0$, then on cursory inspection, equation (6) seems to imply that $p_{i2} = 0$ which is clearly untrue. However, if $z_{i2} = 0$, then π_{i2} must equal 1, since by definition of the conditional probability, individual i has been seen at least once, and $z_{i2} = 0$ means it was not seen on

the first occasion, so it must have been seen on the second occasion. Thus, $\pi_{i2} = 1$ and p_{i2} is ill-defined (the denominator of equation (6) being equal to 0).

By not including past capture histories in their modelling of the conditional probabilities of capture, Buckland *et al.* (1993b) avoid having to exclude all individuals seen only from the second look-out station.

Appendix B: An S-PLUS program to obtain abundance estimates using the logistic regression methodology

The following data are expected to be found in a file called 'datafile':

- SEEN: the observations, y_{ij} .
- ZIJ: the capture histories, z_{ij} .
- AGE: a factor denoting the age category of the individuals (assumed constant throughout the study period).

The following code assumes that the number of sampling occasions (here, 'TRAPOCC') is 5. It calls two customized functions, `get.off` and `get.p` to evaluate the values of the offset and the unconditional capture probabilities, p_{ij} , respectively. The structures of these two functions are outlined at the end of the main program, but without the nuts and bolts of the algebra.

```
function (data = datafile, eps = 0.01)
{
# Main program
  options(contrasts = c("contr.treatment", "contr.poly"))
  numind ← (length(data$SEEN))/5
  TRAPOCC ← rep(1:5, numind)
  newdata ← data.frame(data, TRAPOCC)
  fit1 ← glm(SEEN ~ AGE + ZIJ + TRAPOCC, family = binomial, data= newdata)
  TMPPI ← matrix(fitted(fit1), nrow = numind, ncol = 5, byrow = T) #
# note: TMPPI[1, ] = individual i, capture probs on sampling occasions 1 to 5
# and TMPPI[, 1] = all individuals, capture probs on sampling occasion 1
```

```

# Calculate offset values
  OFFSET ← get.off(TMPPI, numind, data)
  discrep ← 10000
  nloops ← 0 #
# Fit model with offset and iterate until offset doesn't change anymore
  while(discrep > eps) {
    nloops ← nloops + 1
    newdata2 ← data.frame(data, TRAPOCC, OFFSET)
    fit2 ← glm(SEEN ~ AGE + ZIJ + TRAPOCC + offset(OFFSET), family
= binomial, data = newdata2)
    OLDOFF ← OFFSET #
# Calculate new offset values
    TMPPI ← matrix(fitted(fit2), nrow = numind, ncol = 5, byrow = T)
    OFFSET ← get.off(TMPPI, numind, data)
    discrep ← sum((OFFSET - OLDOFF)^2)
    print(cbind(nloops, round(discrep, 3)))
  }
# Fit model with offset converged upon above, then allow stepwise covariate selection
  newdata2 ← data.frame(data, TRAPOCC, OFFSET)
  assign("newdata2", newdata2, frame = 1)
  fit2 ← glm(SEEN ~ AGE + ZIJ + TRAPOCC + offset(OFFSET), family = binomial,
data = newdata2)
  fit.final ← step.glm(fit2, direction = "both") #
# Calculate phats
  TMPPI ← matrix(fitted(fit.final), nrow = numind, ncol = 5, byrow = T)
  PHAT ← get.p(TMPPI, numind, data)
  print(PHAT) #
# To estimate abundance, need prob that each animal was seen at least once
  PROBI ← 1 - ((1 - PHAT[, 1]) * (1 - PHAT[, 2]) * (1 - PHAT[, 3]) * (1 - PHAT[, 4])
* (1 - PHAT[, 5]))
  MHAT ← sum(1/PROBI)
  return(fit.final, numind, MHAT)
}

```

```

function(TMPPI, numind, data)
{ # Function: get.off
  z ← matrix(data$ZIJ, nrow = numind, ncol = 5, byrow = T)
# z1, ..., z5 are vectors of the past capture histories  $z_{i1}, \dots, z_{i5}$ 
  z1 ← z[, 1]
  z2 ← z[, 2]
  z3 ← z[, 3]
  z4 ← z[, 4]
  z5 ← z[, 5]
# c1, ..., c5 are vectors of the conditional capture probabilities  $\pi_{i1}, \dots, \pi_{i5}$ 
  c1 ← TMPPI[, 1]
  c2 ← TMPPI[, 2]
  c3 ← TMPPI[, 3]
  c4 ← TMPPI[, 4]
  c5 ← TMPPI[, 5]
  :
[Algebraic expressions for OFF1, ..., OFF5 in terms of z1, ..., z5 and c1, ..., c5 omitted]
  :
  OFFSET ← NULL
  OFFSET ← c(OFFSET, rbind(OFF1, OFF2, OFF3, OFF4, OFF5))
  return(OFFSET)
}

```

```

function(TMPPI, numind, data)
{ # Function: get.p
  z ← matrix(data$ZIJ, nrow = numind, ncol = 5, byrow = T) #
# z1, ..., z5 are vectors of the past capture histories  $z_{i1}, \dots, z_{i5}$ 
  z1 ← z[, 1]
  z2 ← z[, 2]
  z3 ← z[, 3]
  z4 ← z[, 4]
  z5 ← z[, 5]

```

```

# c1, ..., c5 are vectors of the conditional capture probabilities  $\pi_{i1}, \dots, \pi_{i5}$ 
c1 ← TMPPI[, 1]
c2 ← TMPPI[, 2]
c3 ← TMPPI[, 3]
c4 ← TMPPI[, 4]
c5 ← TMPPI[, 5]
  ⋮

[Algebraic expressions for p1, ..., p5 in terms of z1, ..., z5 and c1, ..., c5 omitted]
  ⋮
PHAT ← cbind(p1, p2, p3, p4, p5)
return(PHAT)
}

```

Appendix C: Maple commands for use in implementing the logistic regression methodology

The following Maple commands can be used to obtain the expressions for the offset term and the unconditional capture probabilities needed by the S-PLUS functions `get.off` and `get.p` outlined in Appendix B. The commands below provide expressions for the case of five sampling occasions.

```

> eq1:=c1=p1/(1-(1-p1)*(1-p2)*(1-p3)*(1-p4)*(1-p5));
> eq2:=c2=p2/(1-(1-z2)*(1-p2)*(1-p3)*(1-p4)*(1-p5));
> eq3:=c3=p3/(1-(1-z3)*(1-p3)*(1-p4)*(1-p5));
> eq4:=c4=p4/(1-(1-z4)*(1-p4)*(1-p5));
> eq5:=c5=p5/(1-(1-z5)*(1-p5));
> eq11:=g1=log(p1/(1-p1));
> eq12:=g2=log(p2/(1-p2));
> eq13:=g3=log(p3/(1-p3));
> eq14:=g4=log(p4/(1-p4));
> eq15:=g5=log(p5/(1-p5));
> solve({eq1,eq2,eq3,eq4,eq5,eq11,eq12,eq13,eq14,eq15},{p1,p2,p3,p4,p5,g1,g2,g3,g4,g5})

```

where

$$c_1, \dots, c_5 = \pi_{i_1}, \dots, \pi_{i_5};$$

$$p_1, \dots, p_5 = p_{i_1}, \dots, p_{i_5}; \text{ and}$$

$$g_1, \dots, g_5 = \ln[\pi_{i_1}/(1 - \pi_{i_1})] + \text{off}_{i_1}, \dots, \ln[\pi_{i_5}/(1 - \pi_{i_5})] + \text{off}_{i_5}.$$

Chapter 7

Summary and Conclusions

The primary aim of this thesis has been the development of methodology that will be of practical use for surveying cetaceans. For the most part (chapters 3-5), the research has focussed on developing methods which allow heterogeneity in the spatial distribution of cetaceans to be modelled using standard line transect data. Two models were proposed. The first of these was based on dividing the transects into smaller units, and modelling the expected number of sightings in each small unit. This model was applied to shipboard line transect data from an Antarctic minke whale survey, and to cue count data from an aerial minke whale survey off Greenland, providing informative results. The second model, which was derived within a point process modelling framework, was based on modelling the waiting interval between successive detections. A method for implementing the model within a GAM framework was suggested — this was also applied to the Antarctic minke whale data. It is particularly encouraging that these methods have already been used by some researchers (eg. Clarke *et al.* (1998; 1999; 2000), Burt *et al.* (1999a; 1999b)), whilst others have shown interest in using them (personal communications from P.S. Hammond; S. Lens; H. Murase; S.B. Reilly). In chapter 6, attention was turned to mark-recapture surveys and a method was developed which allowed heterogeneity in individuals' capture probabilities to be modelled. This was applied to photo-identification data from a study of bottlenose dolphins in the Moray Firth, Scotland, with promising results.

In the remainder of this chapter, the main developments of the thesis are discussed further, noting current research and suggesting areas where future research might be directed.

1 Spatial modelling

As far as we are aware, the only two previous attempts to obtain maps of the spatial distribution of cetaceans from line transect surveys were the contouring and universal kriging approaches taken by Butterworth and Best (1982). The contouring approach was extremely unsatisfactory, being unreproducible and unable to provide variance estimates. Although Butterworth and Best themselves noted these shortcomings, a critical appraisal of the method (de la Mare and Cooke, 1982) duly confirmed that this method should be abandoned in favour of other more feasible analytical techniques. The universal kriging approach was never successfully implemented either. The results from Butterworth and Best (1982) suggest a biologically unrealistic correlation structure, as well as negatively biased variance estimates, and thus the approach was not pursued further. Following these early attempts, there then followed a hiatus in spatial modelling of the distribution of cetaceans, even though the use of spatial models for explaining the distribution of other fauna was becoming more and more widespread (eg. Walker (1990), Osborne and Tigar (1992), Buckland and Elston (1993), Augustin *et al.* (1996a; 1996b)).

In terms of relating cetacean distribution and habitat, substantial progress was made using canonical correspondence analysis (CCA; ter Braak (1986)), for example by Fiedler and Reilly (1994), and Reilly and Fiedler (1994), in modelling the distribution of dolphins in the eastern tropical Pacific. However, whilst CCA is a useful graphical technique, its modelling limitations are discussed in Fewster (1999a). Most notably, it is perhaps most useful for generating hypotheses about relationships, rather than testing these hypotheses. CCA working as well as it possibly can (ie. with 100% of the variance explained) corresponds to the species responses being a linear function of the explanatory variables. Spatial models offer substantially more: different link functions and error structures; GAMs, in which explanatory variables can be incorporated via a smooth flexible function; mixed models; and diagnostics to examine goodness-of-fit. The models developed in this thesis should therefore significantly improve our ability to relate the distribution of cetaceans to their habitat — providing a potentially important means to detect changes in their distribution caused by impacts on their environment, such as the melting of the Antarctic ice shelf induced by global warming.

The primary motivation for this thesis was to develop spatial modelling methodology

capable of estimating abundance by small geographic regions that do not necessarily correspond with survey strata, ie. IWC Small Management Areas (SMAs). The capability to do this, once a suitable model framework was in place, largely follows simply as a result of differences between the adoption of a model-based, rather than a design-based philosophy. Recall that in conventional line transect estimation, it is assumed that the transect lines are placed according to some random (or at least, systematic) sampling design, independently of the distribution of the species under study. With a model-based strategy, this assumption can be relaxed, and data from non-random tracklines may be used in spatial models. This potentially allows estimation from so-called ‘Platforms of Opportunity’, at least for relative abundance (Bravington, 2000b). In the Greenland minke whale example of chapter 3, prediction from the model was restricted to those (mostly coastal) survey blocks with sightings. Of course, given a spatial model, prediction is possible at all values of the explanatory variables, and were this to be done for the remaining offshore blocks, a region of spuriously higher than expected density is predicted to occur within some of these blocks. This example serves to remind us that increased predictive capability is not necessarily an advantage of a model-based approach. Since there is no guarantee that the model will provide a good fit beyond the range of the data (here, there was little effort as well as no sightings in those offshore blocks), extrapolation is not to be recommended in general. In contrast, interpolation is perfectly reasonable, and the ability of spatial models to take advantage of data from outwith some small area, such as a SMA, seems likely to provide increased precision (particularly when there are few sightings in that area).

Whilst potentially useful methodology has been developed in this thesis, there are two main areas of future research that would help provide further insight into their utility for stock management purposes. The first of these is the development of formal model selection criteria for GAMs, including smoothing parameter selection. The second is the incorporation of correlation structure in the spatial models, which should then yield unbiased, or at least less biased, variance estimates. With regard to model selection, Clarke *et al.* (2000) conducted a comparison of the performance of line transect estimates, the count model of chapter 3 and the waiting area model of chapters 4 and 5 (implemented using standard GAM software as proposed in chapter 4), in a simulation study. Their particular interests were in obtaining unbiased estimates from the non-standard survey design used in JARPA (Japanese Research under Special Permit in the Antarctic) surveys, and in investigating how any bias resulting from the spatial models was affected by

the degrees of freedom of the smooth. No comparison of variance estimates was made. In these simulations, the count model performed better than the waiting areas model; for both models, as the degrees of freedom increased, so did the abundance estimates — the size of this effect depending on the underlying distribution (clustering) of whales.

The difficulty in estimating the true clustering process underlying cetacean distribution within some well-defined area should not be underestimated. Cowling (1998) estimates the parameters of a Neyman-Scott clustering process, applying the results to North Atlantic minke whale data, but these methods require more development before being of significant practical use. An alternative, innovative approach has recently been taken by Bravington (2000a). He develops a means of describing the observed clustering directly, using autoregressive models more commonly used in time series applications. Since these models require Normally-distributed data, the ‘gaps’ (ie. the waiting times) between successive sightings are first transformed to a uniform (0,1) distribution, which then undergo a further transformation to obtain Normality. This approach may be used to generate bootstrap samples from which realistic estimates of variance can be obtained, and whilst it would be appropriate to investigate the method further, it represents a promising development in the application of spatial models to line transect data. Further research might also be directed towards developing the approach taken by Gotway and Stroup (1997), in which the correlation structure (estimated via a semi-variogram) is incorporated into the variance-covariance matrix of a GLM; this approach has not as yet been extended to GAMs.

These two main developments aside, a few relatively minor aspects of the modelling would benefit from further investigation. For example, it would be of interest to investigate the robustness of the count data model to different choices of segment length. This is the kind of exercise that is relatively trivial to conduct using simulation, yet would shed light on the general utility of the method. As in Clarke *et al.* (2000), such a simulation should consider a variety of possible clustering scenarios. The modelling presented in chapter 4 was extensively aimed at developing a ‘waiting distance/area’ framework capable of being implemented using standard software. To this end, an iterative procedure was suggested for dealing with changes in expected encounter rate and other variables as the observer moves from one detection to the next. Convergence problems with this procedure were noted, and an *ad hoc* solution was suggested that appeared to solve the problems — at

least for the Antarctic minke whale data. It is probably a little too optimistic to expect this to be the case in general; some further research into the convergence properties of the iterative procedure may help to elucidate the cause of these problems. It is not unreasonable, however, to expect these problems to be alleviated when the correlation structure is incorporated into the model.

If an adequate solution to the convergence problem cannot be found, the likelihood-based approach based on point process modelling (presented in chapter 5) should provide a viable alternative. Currently, the likelihood has been maximized under simplifying assumptions, with a univariate detection function (based only on perpendicular distances) and spatial variation in density being modelled in the along-trackline direction. It is planned to investigate the use of the EM algorithm and MCMC approaches to fit models which allow full implementation of the method. Such a development would provide the potential to model heterogeneity in density and detectability both along and perpendicular to the tracklines.

2 Mark-recapture modelling

Given that a major development in the field of mark-recapture modelling would be required to incorporate heterogeneity in individuals' capture probabilities for use in open population models, the methodology developed in chapter 6 represents a reasonable intermediate approach, providing an adequate, practical solution to the specific research question that was posed. Although the material is not statistically ground-breaking, it does represent the development of a new method to model closed population mark-recapture data. Whether or not this method is taken up by other scientists remains to be seen — it is certainly easy to implement and our results from applying the method to the Moray Firth bottlenose dolphin photo-identification data indicate quite a substantial improvement in precision. This is perhaps the most encouraging aspect of the methodology, and if improved precision is demonstrated to be the case in general, then it would be useful in increasing the power to detect trend from a series of abundance estimates — a valuable tool in the monitoring of populations considered at risk.

Bibliography

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petran and Csàaki (eds), *Second International Symposium on Information Theory*, pp. 267–81. Akademia Kiadó, Budapest.

Alho, J. M. (1990). Logistic regression in capture-recapture models. *Biometrics*, **46**, 623–35.

Augustin, N. H. (1999). Spatial and spatio-temporal models with applications in vegetation dynamics and wildlife population estimation. Ph.D. thesis, University of St Andrews. 177pp.

Augustin, N. H., Borchers, D., Muggleston, M. A., and Buckland, S. T. (1996a). Regression methods with spatially referenced data. *Aspects of Applied Biology*, **46**, 67–74.

Augustin, N. H., Muggleston, M. A., and Buckland, S. T. (1996b). An autologistic model for the spatial distribution of wildlife. *J. Appl. Ecol.*, **33**, 339–47.

Augustin, N. H., Muggleston, M. A., and Buckland, S. T. (1998). The role of simulation in modelling spatially correlated data. *Environmetrics*, **9**, 175–96.

Barlow, J. (1993). The abundance of cetaceans in california waters estimated from ship-board surveys in summer/fall 1991. SWFSC Administrative report LJ-93-09 presented during the 1993 Status of Cetacean Stocks Workshop (unpublished). 39pp.

Beavers, S. C. and Ramsey, F. L. (1998). Detectability analysis in transect surveys. *J. Wildl. Man*, **62**, 948–57.

Best, P. B. and Butterworth, D. S. (1980). Report of the southern hemisphere minke whale assessment cruise, 1978/79. *Rep. int. Whal. Commn*, **30**, 257–83.

- Borchers, D. L. (1994). Methods of estimating mean school size from IWC sightings survey data. *Rep. int. Whal. Commn*, **44**, 429–37.
- Borchers, D. L., Buckland, S. T., Clarke, E. D., and Cumberworth, S. L. (1995). Estimation of cetacean abundance from the SCANS shipboard survey in summer 1994. Paper SC/47/SM36 presented to the Scientific Committee of the International Whaling Commission, May 1995 (unpublished). 41pp.
- Borchers, D. L., Buckland, S. T., Goedhart, P. W., Clarke, E. D., and Hedley, S. L. (1998). Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics*, **54**, 1221–37.
- Borchers, D. L., Buckland, S. T., Priede, I. G., and Ahmadi, S. (1997). Improving the precision of the daily egg production method using generalized additive models. *Can. J. Fish. Aquat. Sci*, **54**, 2727–42.
- Borchers, D. L. and Cameron, C. (1995). Analysis of the 1992/93 IWC minke whale sightings survey in area III. Paper SC/47/SH16 presented to the Scientific Committee of the International Whaling Commission, May 1995 (unpublished). 16pp.
- Borchers, D. L. and Chalis, S. (1996). An overview of the IWC Database-Estimation Software System (DESS). Paper SC/48/O36 presented to the Scientific Committee of the International Whaling Commission, June 1996 (unpublished). 4pp.
- Bravington, M. V. (2000a). Coping with clustered sightings. Paper SC/52/RMP15 presented to the Scientific Committee of the International Whaling Commission, June 2000 (unpublished). 10pp.
- Bravington, M. V. (2000b). Covariate models for continuous-time sightings data. Paper SC/52/RMP14 presented to the Scientific Committee of the International Whaling Commission, June 2000 (unpublished). 15pp.
- Buckland, S. T. (1987). An assessment of the performance of line transect models for fitting IWC/IDCR cruise data, 1978/79 to 1984/85. *Rep. int. Whal. Commn*, **37**, 277–79.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., and Laake, J. L. (1993a). *Distance Sampling: estimating abundance of biological populations*. Chapman and Hall, London.

446pp.

Buckland, S. T. and Anganuzzi, A. A. (1988). Comparison of smearing methods in the analysis of minke sightings data from IWC/IDCR Antarctic cruises.. *Rep. int. Whal. Commn*, **38**, 257–63.

Buckland, S. T., Breiwick, J. M., Cattanach, K. L., and Laake, J. L. (1993b). Estimated population size of the California gray whale. *Mar. Mamm. Sci*, **9**, 235–49.

Buckland, S. T. and Elston, D. A. (1993). Empirical models for the spatial distribution of wildlife. *J. Appl. Ecol*, **30**, 478–95.

Buckland, S. T. and Turnock, B. J. (1992). A robust line transect method. *Biometrics*, **48**, 901–9.

Burnham, K. P., Anderson, D. R., and Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. *Wildl. Monog*, **72**.

Burt, M. L., Buckland, S. T., and Hedley, S. L. (1999a). Spatial modelling of whale and dolphin distribution in the western Mediterranean. Internal RUWPA report, June 1999.

Burt, M. L., Hedley, S. L., Borchers, D. L., and Buckland, S. T. (1999b). Spatial modelling of data from Project SCANS. Internal RUWPA report, January 1999.

Butterworth, D. S. (1982). On the functional form used for $g(y)$ for minke whale sightings, and bias in its estimation due to measurement inaccuracies. *Rep. int. Whal. Commn*, **32**, 883–88.

Butterworth, D. S. (1988). Report of the sub-committee on Southern Hemisphere minke whales, Appendix 4. An alternative method of incorporating school size information into Antarctic minke sighting abundance estimates. *Rep. int. Whal. Commn*, **38**, 84.

Butterworth, D. S. (1991). Report of the sub-committee on Southern Hemisphere minke whales, Appendix 4. The use of $g(0) = 1$ for Southern Hemisphere minke whale sighting surveys.. *Rep. int. Whal. Commn*, **41**, 126.

Butterworth, D. S. and Best, P. (1982). Report of the southern hemisphere minke whale assessment cruise, 1980/81. *Rep. int. Whal. Commn*, **32**, 835–69.

Butterworth, D. S., Best, P. B., and Basson, M. (1982). Results of analysis of sighting experiments carried out during the 1980/81 Southern Hemisphere minke whale assessment cruise. *Rep. int. Whal. Commn*, **32**, 819–34.

Butterworth, D. S., Best, P. B., and Hembree, D. (1984). Analysis of experiments carried out during the 1981/82 IWC/IDCR Antarctic minke whale assessment cruise in Area II. *Rep. int. Whal. Commn*, **34**, 365–92.

Butterworth, D. S. and Borchers, D. L. (1988). Estimates of $g(0)$ for minke schools from the results of the independent observer experiment on the 1985/86 and 1986/87 IWC/IDCR Antarctic assessment cruises. *Rep. int. Whal. Commn*, **38**, 301–13.

Butterworth, D. S., Haw, M. D., and Borchers, D. L. (1992). Report on progress in developing estimates of abundance for Southern Hemisphere minke whales by 5° longitude slice. Paper SC/44/SHB3 (and Addendum) presented to the Scientific Committee of the International Whaling Commission, June 1992 (unpublished). 17pp.

Chambers, J. M. and Hastie, T. J. (1993). *Statistical models in S*. Chapman and Hall, New York. 608pp.

Chen, S. (1994). Studying school size effects in line transect sampling using the kernel method. Research Report, CSIRO Division of Fisheries, Hobart.

Chen, S. X. (1998). Measurement errors in line transect surveys. *Biometrics*, **54**.

Chen, S. X. and Polacheck, T. (1996). Kernel estimates of mean school size for minke whales from IWC sighting surveys. *Rep. int. Whal. Commn*, **46**, 341–48.

Clarke, E. D., Ashbridge, J., Burt, M. L., Hedley, S. L., and Borchers, D. L. (1998). GAM-based abundance estimation from JARPA survey data: progress and simulation model design. Paper SC/50/CAWS33 presented to the Scientific Committee of the International Whaling Commission, May 1998 (unpublished). 51pp.

Clarke, E. D., Burt, M. L., and Borchers, D. L. (1999). Simulation of JARPA surveys to test abundance estimation methods. Paper SC/51/RMP16 presented to the Scientific Committee of the International Whaling Commission, May 1999 (unpublished). 29pp.

Clarke, E. D., Burt, M. L., and Borchers, D. L. (2000). Investigation of bias of GAM-

based abundance estimation methods and their suitability for JARPA survey data. Paper SC/52/IA19 presented to the Scientific Committee of the International Whaling Commission, June 2000 (unpublished). 15pp.

Cooke, J. G. (1985). Notes on the estimation of whale density from line transects. *Rep. int. Whal. Commn*, **35**, 319–23.

Cooke, J. G. (1996). A modification of SC/47/NA14 implementation of the hazard-rate method to accommodate covariates. Paper SC/A96/AE24 presented to the IWC Abundance Assessment Working Group Meeting, Oslo, April-May 1996 (unpublished). 11pp.

Cooke, J. G. (1997). An implementation of a surfacing-based approach to abundance estimation of minke whales from shipborne surveys. *Rep. int. Whal. Commn*, **47**, 513–28.

Cooke, J. G. and Leaper, R. (1998). A general modelling framework for the estimation of whale abundance from line transect surveys. Paper SC/50/RMP21 presented to the Scientific Committee of the International Whaling Commission, June 1998 (unpublished). 16pp.

Cowling, A. (1998). Spatial methods for line transect surveys. *Biometrics*, **54**, 828–39.

Cox, D. R. and Isham, V. (1980). *Point Processes*. Chapman and Hall, London. 188pp.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions. *Numer. Math*, **31**, 377–403.

Cressie, N. A. C. (1991). *Statistics for Spatial Data*. John Wiley and Sons, New York. 900pp.

Curran, S., Wilson, B., and Thompson, P. M. (1996). Recommendations for sustainable management of the bottlenose dolphin population in the Moray Firth. Scottish Natural Heritage Review Number 56. Battleby, UK.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press, Cambridge. 582pp.

de la Mare, W. K. and Cooke, J. G. (1982). A review of the stratification of sightings data used in the estimation of populations of Southern Hemisphere minke whales (*balaenoptera*

- acutorostrata*). *Rep. int. Whal. Commn*, **32**, 315–322.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B*, **39**, 1–38.
- Doi, T. (1971). Further development of sighting theory on whales. *Bull. Tokai Reg. Fish. Res. Lab*, **68**, 1–22.
- Doi, T. (1974). Further development of whale sighting theory. In W. Schevill (ed.), *The Whale Problem: a Status Report*, pp. 359–68. Harvard University Press, Cambridge, Massachusetts.
- Doi, T., Kasamatsu, F., and Nakano, T. (1982). A simulation study on sighting survey of minke whales in the Antarctic. *Rep. int. Whal. Commn*, **32**, 919–28.
- Doi, T., Kasamatsu, F., and Nakano, T. (1983). Further simulation studies on sighting by introducing both concentration of sighting effort by angle and aggregations of minke whales in the Antarctic. *Rep. int. Whal. Commn*, **33**, 403–12.
- Drummer, T. D. and McDonald, L. L. (1987). Size bias in line transect sampling. *Biometrics*, **43**, 13–21.
- Efron, B. and Tibshirani, R. J. (1993). *Monographs on Statistics and Applied Probability*. 57. *An introduction to the bootstrap*. Chapman and Hall, New York. 436pp.
- Fancy, S. G. (1997). A new approach for analyzing bird densities from variable circular plot counts. *Pacific Science*, **51**, 107–14.
- Fewster, R. M. (1999a). Overview of multivariate ordination methods for ecological data. Paper SC/M99/SOWER11 presented to the IWC Scientific Committee SOWER 2000 Workshop, March 1999, Edinburgh (unpublished). 14pp.
- Fewster, R. M. (1999b). Population dynamics models in a spatial oceanographic setting. Paper SC/M99/SOWER2 presented to the IWC Scientific Committee SOWER 2000 Workshop, March 1999, Edinburgh (unpublished). 20pp.
- Fiedler, P. C. and Reilly, S. B. (1994). Interannual variability of dolphin habitats in the eastern tropical Pacific. 2. Effects on abundances estimated from tuna vessel sightings,

1975-1990. *Fish. Bull.*, **92**, 451-63.

Gamerman, D. (1997). *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference*. Chapman and Hall, London. 245pp.

Gotway, C. A. and Stroup, W. W. (1997). A generalized linear model approach to spatial data analysis and prediction. *J. Agric. Biol. and Environ. Statist.*, **2**, 157-78.

Hammond, P. S., Benke, H., Berggren, P., Borchers, D. L., Buckland, S. T., Collet, A., Heide-Jørgensen, M. P., Heimlich-Boran, S., Hiby, A. R., Leopold, M. P., and Øien, N. (1995). Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters. *LIFE 92-2/UK/027 Final Report.*, 242pp.

Hastie, T. J. and Tibshirani, R. J. (1990). *Monographs on Statistics and Applied Probability. 43. Generalized Additive Models*. Chapman and Hall, London. 335pp.

Hedley, S. L., Buckland, S. T., and Borchers, D. L. (1997a). A review of analysis options for implementation in the IWC Database and Estimation Software System. *Rep. int. Whal. Commn.*, **47**, 885-95.

Hedley, S. L., Buckland, S. T., and Borchers, D. L. (1999). Spatial modelling from line transect data. *J. Cetacean Res. Manage.*, **1**, 255-64.

Hedley, S. L., Neve, P. B., and Borchers, D. L. (1997b). Abundance of minke whales off West Greenland, 1993. Paper SC/49/NA7 presented to the Scientific Committee of the International Whaling Commission, October 1997 (unpublished). 11pp.

Hiby, A. R. and Hammond, P. S. (1989). Survey techniques for estimating current abundance and monitoring trends in abundances of cetaceans. *Rep. int. Whal. Commn (special issue 11)*, pp. 47-80.

Hiby, A. R. and Lovell, P. (1998). Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics*, **54**, 1280-89.

Hiby, A. R., Ward, A., and Lovell, P. (1989). Analysis of the North Atlantic Sightings Survey 1987: final aerial survey results. *Rep. int. Whal. Commn.*, **39**, 447-55.

Högmander, H. (1991). A random field approach to transect counts of wildlife populations.

Biom. J., **33**, 1013–23.

Högmander, H. (1995). *Methods of spatial statistics in monitoring wildlife populations*, Jyväskylä Studies in Computer Science, Economics and Statistics 25. University of Jyväskylä. 98pp.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.*, **47**, 663–85.

Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika*, **76**, 133–40.

Huggins, R. M. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*, **47**, 725–32.

IWC (1997). Report of the Abundance Estimation Working Group to the IWC Scientific Committee. *Rep. int. Whal. Commn.*, **47**, 261–90.

IWC (1999). Report of the Scientific Committee. Annex N. The Revised Management Procedure (RMP) for baleen whales. *J. Cetacean Res. Manage. (Suppl.)*, **1**, 251–8.

Janik, V. M. and Thompson, P. M. (1996). Changes in the surfacing patterns of bottlenose dolphin in response to boat traffic. *Mar. Mamm. Sci.*, **12**, 597–601.

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration — stochastic model. *Biometrika*, **52**, 225–47.

Kayes, R. (1985). The decline of porpoises and dolphins in the southern North Sea: a current status report. Number 14. Political Ecology Research Group, Oxford, UK.

Kendall, W. L. and Pollock, K. H. (1992). The robust design in capture-recapture studies: a review and evaluation by Monte Carlo simulation. In D. R. McCullough and R. H. Barrett (eds), *Wildlife 2001: Populations*, pp. 31–43. Elsevier, London.

Kendall, W. L., Pollock, K. H., and Brownie, C. (1995). A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics*, **51**, 293–308.

- Kishino, H., Kasamatsu, F., and Miyashita, T. (1986). Report of the Workshop on Minke Whale Sightings, annex f. difference of detection function between two ships in PS experiment and eh factor. *Rep. int. Whal. Commn*, **36**, 467–68.
- Laake, J. L., Buckland, S. T., Anderson, D. R., and Burnham, K. P. (1996). *DISTANCE 2.2*. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO. Available: <http://www.ruwpa.st-and.ac.uk/distance/>.
- Larsen, F. (1995). Abundance of minke and fin whales off West Greenland, 1993. *Rep. int. Whal. Commn*, **45**, 365–70.
- Marques, F. F. C. (in prep.). Abundance estimation from opportunistic line transect surveys: an application to tuna vessel observer data. Ph.D. thesis, University of St Andrews.
- Matheron, G. (1969). *Le Krigeage Universel*. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau. (In French).
- McCullagh, P. and Nelder, J. A. (1989). *Monographs on Statistics and Applied Probability. 37. Generalized Linear Models*, 2nd edition. Chapman and Hall, London. 511pp.
- Miller, R. G. (1974). The jackknife — a review. *Biometrika*, **61**, 1–17.
- Müller, H. G. (1987). Weighted local regression and kernel methods for nonparametric curve fitting. *J. Am. Statist. Assoc*, **82**, 231–8.
- Nelder, J. A. and Mead, R. (1965). *Computer Journal*, **7**, 308.
- Osborne, P. E. and Tigar, B. J. (1992). Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. *J. Appl. Ecol*, **29**, 55–62.
- Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildl. Monog*, **62**.
- Palka, D. L. (1993). Estimating density of animals when assumptions of line transect survey are violated. Ph.D. thesis, University of California, San Diego. 169pp.
- Polacheck, T. (1995). The effect of increasing observer trackline effort in shipboard line transect surveys for harbor porpoise. *Rep. int. Whal. Commn (special issue 16)*, pp. 69–88.

- Pollock, K. H. (1982). A capture-recapture design robust to unequal probability of capture. *J. Wildl. Man*, **46**, 752–7.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986). *Numerical Recipes*. Press Syndicate of the University of Cambridge, Cambridge. 818pp.
- Quang, P. X. (1991). A nonparametric approach to size-biased line transect sampling. *Biometrics*, **47**, 269–79.
- Quinn, T. J. (1985). Line transect estimators for schooling populations. *Fisheries Research*, **3**, 183–99.
- Ramsey, F. L., Wildman, V., and Engbring, J. (1987). Covariate adjustments to effective area in variable-area wildlife surveys. *Biometrics*, **43**, 1–11.
- Reilly, S. B. and Fiedler, P. C. (1994). Interannual variability of dolphin habitats in the eastern tropical Pacific. 1. Research vessel surveys, 1986–1990. *Fish. Bull.*, **92**, 434–50.
- Schweder, T. (1974). Transformations of point processes: applications to animal sighting and catch problems, with special emphasis on whales. Ph.D. thesis, University of California, Berkeley. 183pp.
- Schweder, T. (1977). Point process models for line transect experiments. In J. R. Barba, F. Brodeau, G. Romier, and B. Van Cutsem (eds), *Recent Developments in Statistics*, pp. 221–43. North Holland Publishing Company, New York.
- Schweder, T. (1990). Independent observer experiments to estimate the detection function in line transect surveys of whales. *Rep. int. Whal. Commn*, **40**, 349–55.
- Schweder, T. (1996). A note on a buoy-sighting experiment in the North Sea in 1990. *Rep. int. Whal. Commn*, **46**, 383–85.
- Schweder, T. (1997). Measurement error models for the Norwegian minke whale survey in 1995. *Rep. int. Whal. Commn*, **47**, 485–488.
- Schweder, T., Øien, N., and Høst, G. (1991). Estimates of the detection probability for shipboard surveys of northeastern Atlantic minke whales, based on a parallel ship experiment. *Rep. int. Whal. Commn*, **41**, 417–32.

- Schweder, T., Øien, N., and Høst, G. (1993). Estimates of abundance of northeastern Atlantic minke whales in 1989. *Rep. int. Whal. Commn*, **43**, 323–31.
- Schweder, T., Skaug, H. J., Dimakos, X. K., Langaas, M., and Øien, N. (1997). Abundance of northeastern Atlantic minke whales, estimates for 1989 and 1995. *Rep. int. Whal. Commn*, **47**, 453–83.
- Scott, D. W. (1992). *Multivariate density estimation*. Wiley, New York. 332pp.
- Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika*, **52**, 249–59.
- Seber, G. A. F. (1982). *The estimation of animal abundance and related parameters*, 2nd edition. Edward Arnold, London. 654pp.
- Silverman, B. W. (1984). A fast and efficient cross-validation method for smoothing parameter choice in spline regression. *J. Amer. Statist. Assoc.*, **79**, 584–9.
- Skaug, H. J. (1997). Perpendicular distance line transect methods based on integrated spatial hazard probability models. *Rep. int. Whal. Commn*, **47**, 493–97.
- Skaug, H. J. and Schweder, T. (1999). Hazard models for line transect surveys with independent observers. *Biometrics*, **55**.
- Stoyan, D. (1982). A remark on the line transect method. *Biom. J.*, **24**, 191–5.
- Strindberg, S. (in prep.). Automated optimised survey design. Ph.D. thesis, University of St Andrews.
- Sullivan, P. J. (1992). A Kalman Filter approach to catch-at-length analysis. *Biometrics*, **48**, 237–57.
- ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167–79.
- Thomas, L., Laake, J. L., Derry, J. F., Buckland, S. T., Borchers, D. L., Anderson, D. R., Burnham, K. P., Strindberg, S., Hedley, S. L., Burt, M. L., Marques, F. F. C., Pollard, J. H., and Fewster, R. M. (1998). *Distance 3.5*. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. Available: <http://www.ruwpa.st->

and.ac.uk/distance/.

Thompson, S. K. (1992). *Sampling*. Wiley, New York. 343pp.

Trenkel, V. M., Elston, D. A., and Buckland, S. T. (2000). Fitting population dynamics models to count and cull data using sequential importance sampling. *J. Amer. Statist. Assoc.*, **95**, 363–74.

Walker, P. A. (1990). Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. *J. Biogeogr.*, **17**, 279–89.

Waterloo Maple Inc. (1994). *Maple V (Release 3)*. Waterloo, Ontario, Canada.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, **61**, 439–47.

Wells, R. S. and Scott, M. D. (1990). Estimating bottlenose dolphin parameters from individual identification and capture-recapture techniques. *Rep. int. Whal. Commn (special issue 12)*, pp. 407–15.

Wilson, B. (1995). The ecology of bottlenose dolphins in the Moray Firth, Scotland: a population at the northern extreme of the species' range. Ph.D. thesis, University of Aberdeen.

Wilson, B., Thompson, P. M., and Hammond, P. S. (1997a). Habitat use by bottlenose dolphins: seasonal distribution and stratified movement patterns in the Moray Firth, Scotland. *J. Appl. Ecol.*, **34**, 1365–74.

Wilson, B., Thompson, P. M., and Hammond, P. S. (1997b). Skin lesions and physical deformities in bottlenose dolphins in the Moray Firth: population prevalence and age-sex differences. *Ambio*, **26**, 243–247.

Wilson, B., Thompson, P. M., and Hammond, P. S. (1999). Estimating size and assessing trends in a coastal bottlenose dolphin population. *Ecological Applications*, **9**, 288–300.

Wood, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *J. R. Statist. Soc. B*, **62**, 413–28.

Zahl, S. (1989). Line transect sampling with unknown probability of detection along the transect. *Biometrics*, **45**, 453–70.