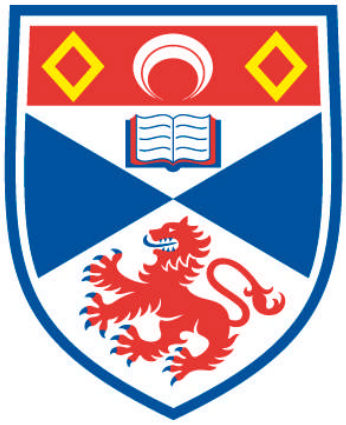
ESTIMATING WILDLIFE DISTRIBUTION AND ABUNDANCE FROM LINE TRANSECT SURVEYS CONDUCTED FROM PLATFORMS OF OPPORTUNITY

Fernanda F. C. Marques

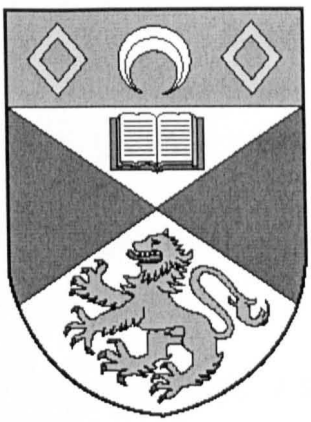
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2001

Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity

Fernanda F. C. Marques



Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

in the School of Mathematics and Statistics,

UNIVERSITY OF ST ANDREWS.

May, 2001.

Declarations

1. I, Fernanda F. C. Marques, hereby certify that this thesis, which is approximately 26,000. 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 ..6/6/01.. signature of candidate

2. I was admitted as a research student in May 1996 and as a candidate for the degree of PhD in May 1997; the higher study for which this is a record was carried out in the University of St Andrews between 1996 and 2001.

   date ..6/6/01.. 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 ..13/6/01.. 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 ..6/6/01.. signature of candidate

Acknowledgments

I would like to thank my supervisor, Prof. Steve Buckland, for his patience, help and support, without which this thesis would have never been possible. I would also like to thank the staff of the Research Unit for Wildlife Population Assessment (RUWPA) at the University of St Andrews, especially Drs. David Borchers, Louise Burt, Liz Clarke and Len Thomas, for sharing their knowledge of statistics and computing with me. Dr. Simon Wood always found the time to answer my questions, for which I am grateful. I am also very grateful to Sam Strindberg, Sharon Hedley, Rachel Fewster, Miguel Bernal and Nicole Augustin, for their help, encouragement, and friendship. Finally, I would like to thank my parents for their unconditional support throughout.

Funding for this work was provided by the Inter-American Tropical Tuna Commission (IATTC), for which I am very thankful. Additional financial support provided by an Overseas Research Student (ORS) Award is gratefully acknowledged.

Abstract

Line transect data obtained from 'platforms of opportunity' are useful for the monitoring of long term trends in dolphin populations which occur over vast areas, yet analyses of such data are problematic due to violation of fundamental assumptions of line transect methodology. In this thesis we develop methods which allow estimates of dolphin relative abundance to be obtained when certain assumptions of line transect sampling are violated.

Generalised additive models are used to model encounter rate and mean school size as a function of spatially and temporally referenced covariates. The estimated relationship between the response and the environmental and locational covariates is then used to obtain a predicted surface for the response over the entire survey region. Given those predicted surfaces, a density surface can then be obtained and an estimate of abundance computed by numerically integrating over the entire survey region. This approach is particularly useful when search effort is not random, in which case standard line transect methods would yield biased estimates.

Estimates of $f(0)$ (the inverse of the effective strip (half-)width), an essential component of the line transect estimator, may also be biased due to heterogeneity in detection probabilities. We developed a conditional likelihood approach in which covariate effects are directly incorporated into the estimation procedure. Simulation results indicated that the method performs well in the presence of size-bias. When multiple covariates are used, it is important that covariate selection be carried out.

As an example we applied the methods described above to eastern tropical Pacific dolphin stocks. However, uncertainty in stock identification has never been directly incorporated into methods used to obtain estimates of relative or absolute abundance. Therefore we illustrate an approach in which trends in dolphin relative abundance are monitored by small areas, rather than stocks.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Knowledge of population size plays a crucial role in wildlife ecology and environmental biology. Estimates of relative or absolute abundance are used to monitor the status of vulnerable and endangered species, to test hypotheses concerning the mechanisms dictating population dynamics and trends, and to devise management strategies for particular ecosystems and/or populations, among other applications.

For many populations, particularly of marine species, complete censuses are not feasible, and abundance estimates are based on sampling methods. Seber (1982) provides a comprehensive review of sampling-based abundance estimation procedures. The methods he describes can be classified into two general categories: capture-recapture methods, in which abundance estimates are obtained based on ratios of the number of individuals either marked (mark-recapture) or removed (change in ratio, catch-effort) from the population at a first capture occasion to the number recaptured/sampled at a second occasion; and sightings survey methods (Buckland *et al.* 1993), where animal density is estimated either by counting the number of animals located within survey strips (strip transects), or by modelling the radial distance of sightings from survey points (point transect sampling) or the perpendicular distance of sightings from the transect line (line transect sampling). Of these, line transect sampling (Seber 1982, Buckland *et al.* 1993) is the most commonly used method for populations of marine mammals which occur over vast areas.

Unlike capture-recapture methods, line transect sampling does not require repeated surveys in order to yield an estimate of abundance. In addition, when the population of interest is large, sample sizes required by other methods to yield abundance estimates with a level of precision compatible with the required objectives can be prohibitive. Take the Petersen method as an example, in which abundance estimates are based on the

product of the number of animals captured at the first and second sampling occasions, respectively, divided by the number of animals captured at both occasions (this corresponds to the method originally proposed, but a modified version of the method is usually applied instead; for details see Chapman (1951) and Seber (1982)). Assume that individuals can be relatively easily marked. Following Seber (1982, p.64), let the accuracy of the estimator be defined as the margin by which the estimated population size may differ from the true population size, at 95% confidence levels. Then, given a true population of 500 000 individuals, the Petersen method would require approximately 0.7% of the population to be sampled at each capture occasion to yield an estimate of the true population size within 50% of the true population abundance (Appendix 1.5.1). Although this percentage is low, it corresponds to about 3 500 individuals being marked/recaptured at each sampling occasion, a non-negligible number considering that the population may be spread over a large area. If more precise estimates of abundance are required, say within 25% of the true population abundance (at the 95% confidence level), the number of individuals to be captured at each occasion rises to approximately 5 850. In contrast, abundance estimates for populations of comparable size based on line transect data can be obtained using sample sizes of less than 400. As an example we estimate the expected number of detections under two different required levels of precision using published data from dolphin sightings surveys (Appendix 1.5.2). For a target coefficient of variation (CV) of 50%, a total of 94 sightings are estimated to be made, whereas for a CV of 25% the estimated number of detections rises to 378. These results are consistent with examples from the literature (e.g. Forney and Barlow 1993).

This thesis addresses methodological issues associated with line transect data when certain assumptions of the method are violated.

1.1 Line transect sampling

The derivation of the line transect estimator presented below is taken from Seber (1982) and Buckland *et al.* (1993). Given a region of size A containing N objects, let there be k transect lines of total length L ($L = \sum_{j=1}^{k} L_j$) randomly or systematically placed over the region so that each point in the region has equal probability of being sampled (Figure 1.1). An observer travels along each transect and records the perpendicular distance x_i ($i = 1, \ldots, n$) of each detected object from the transect line. In many cases it is not feasible to measure the perpendicular distances directly, so instead the radial distance r_i and sighting angle δ_i to each detected object are recorded and later converted to perpendicular distance.

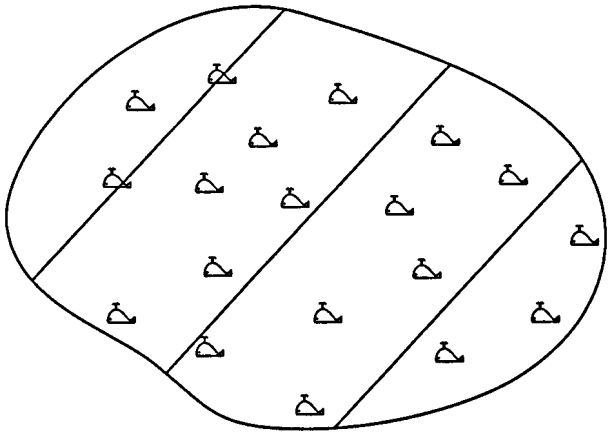


Figure 1.1: Survey region showing three transect lines placed in a systematic fashion. Objects in this case correspond to whales.

Although objects may be detected out to ∞, for simplicity we assume that only objects out to some distance W from the transect line are recorded, so that $0 \leq x_i \leq W$. Denote by n_j the number of objects detected in transect j within distance W from the transect line, and let $n = \sum_{j=1}^{k} n_j$ indicate the total number of detected objects. The abundance of objects in the survey region is given by (Buckland *et al.* 1993):

$$N = A \cdot D, \tag{1.1}$$

where D denotes the density of objects, given by:

$$D = \frac{\mathbb{E}[n]}{2LW \cdot P} \tag{1.2}$$

and estimated by:

$$\hat{D} = \frac{n}{2LW \cdot \hat{P}}. \tag{1.3}$$

Here $2LW$ denotes the area of the survey region which has been surveyed and P, estimated by $\hat{P}$, is the unconditional probability that an object within the surveyed strip is detected. If objects occur in clusters, then an estimate of the expected cluster size, $\mathbb{E}[s]$, must be incorporated into the expression above:

$$\hat{D} = \frac{n \cdot \hat{\mathbb{E}}[s]}{2LW \cdot \hat{P}}. \tag{1.4}$$

Defining $g(x)$ as the probability that an object is detected, given that it is at distance x from the transect line, and assuming that all objects located on the line are detected with certainty, that is, $g(0) = 1$, then P is given by:

$$\begin{aligned} P &= \int_0^W g(x)\,\pi(x)\,dx \\ &= \frac{\int_0^W g(x)\,dx}{W} \\ &= \frac{\mu}{W}, \end{aligned} \tag{1.5}$$

with $\mu = \int_0^W g(x)\,dx$. It can be shown (Seber 1982, Buckland *et al.* 1993) that $\mu = 1/f(0)$, where $f(0)$ denotes the probability density function (pdf) of the *observed* perpendicular distances evaluated at $x = 0$. Hence the estimator of the density of objects in the survey region is given by:

$$\begin{aligned} \hat{D} &= \frac{n}{2LW \cdot \frac{\hat{\mu}}{W}} \\ &= \frac{n \cdot \hat{f}(0)}{2L} \end{aligned} \tag{1.6}$$

and, in the case where objects occur in clusters, the expression above becomes:

$$\hat{D} = \frac{n \cdot \hat{\mathbb{E}}[s] \cdot \hat{f}(0)}{2L}. \tag{1.7}$$

Assuming no correlation between the various terms, the variance of the density estimate can be approximated using the delta method (p.7-9 in Seber 1982), so that:

$$\widehat{var}(\hat{D}) = \hat{D}^2 \cdot \left\{ \frac{\widehat{var}(n)}{n^2} + \frac{\widehat{var}\{\hat{f}(0)\}}{\{\hat{f}(0)\}^2} \right\} \tag{1.8}$$

and, in the case where objects occur in clusters:

$$\widehat{var}(\hat{D}) = \hat{D}^2 \cdot \left\{ \frac{\widehat{var}(n)}{n^2} + \frac{\widehat{var}\{\hat{\mathbb{E}}[s]\}}{\{\hat{\mathbb{E}}[s]\}^2} + \frac{\widehat{var}\{\hat{f}(0)\}}{\{\hat{f}(0)\}^2} \right\}. \tag{1.9}$$

1.2 Assumptions of line transect methodology

The above estimators rely on the following assumptions (Buckland *et al.* 1993):

(i) Transect lines are randomly placed over the survey region according to some survey design, so that all points in the region have equal coverage probability.

(ii) All objects located on the line are detected with certainty (*i.e.* $g(0) = 1$).

(iii) Objects are detected before any responsive movement.

(iv) Non-responsive movement is slow relative to the speed of the observer.

(v) Measurements are recorded without error.

(vi) Detections are independent events.

(vii) Transect lines are randomly placed with respect to the objects.

Line transect methodology is explicitly based on the assumption of equal coverage probabilities throughout the survey region, so that the probability that an object falls within the area surveyed is given by $2LW/A$. In addition, random line placement ensures that all objects within the area surveyed are uniformly distributed in the interval $[0, W]$, that is, the pdf of the perpendicular distances of all objects within distance W from the transect line, $\pi(x)$, is given by $\pi(x) = W^{-1}$. This latter result forms the basis of the definition of P (*cf.* expression (1.5)). Violation of assumption (i) is often caused by non-random sampling, and potentially introduces bias in the density estimates as the area surveyed may not contain densities representative of the true density over the entire survey region.

Assumption (ii) is the most fundamental assumption of line transect sampling. If $g(0) < 1$, then the estimators given by expressions (1.6) and (1.7) will underestimate the true density of objects in the survey region. Although methods to estimate $g(0)$ have been developed (e.g. Butterworth and Borchers 1988, Schweder 1990, Borchers 1996, Borchers *et al.* 1998*a*), they all require two independent platforms simultaneously surveying the region. Thus unless $g(0)$ estimation is explicitly incorporated into the survey design, unbiased estimation based on standard (*i.e.* single platform) line transect methodology can only be achieved by ensuring that all objects located on the line are detected.

Assumptions (iii) and (iv) are only applicable to mobile objects. If objects systematically avoid the observer, and such responsive movement occurs prior to the objects being

detected, then the observed distribution of the perpendicular distances of the detected objects will exhibit a very low number of detections near the transect line, potentially leading to a substantial underestimate of $f(0)$, and hence of density. Conversely, if objects tend to turn towards the observer before they are detected, then there will be a large number of observations near the transect line, which will result in a substantial overestimate of $f(0)$. Bias in $f(0)$ estimates may also result if non-responsive movement is fast relative to the speed of the observer (Hiby 1986). Approaches which are robust to the violation of this assumption are presented in Buckland and Turnock (1992) and Hammond *et al.* (1995).

Violation of assumption (v) is usually caused by rounding of the observed (radial or perpendicular) distances and/or sighting angles, and results in 'heaping' at certain perpendicular distances, which can lead to poor estimates of the detection function. Heaping at zero is particularly problematic (*cf.* the discussion of the effects of violation of assumptions (iii) and (iv) above). Approaches for the modelling of measurement error have been proposed by Chen (1998) and Cooke and Leaper (1998).

Assumptions (vi) and (vii) are required for unbiased estimation of the variances. However, if empirical estimators for the variance are used, then these assumptions may be relaxed.

Two additional assumptions often mentioned are that (viii) objects are not counted more than once and (ix) the detection function $g(x)$ has a 'shoulder' (*i.e.* $g'(0) = 0$). Violation of the former does not necessarily introduce bias in the resulting estimates, unless objects are being systematically counted more than once within the same transect line, which would result in an overestimate of encounter rate, and hence of density. Violation of assumption (ix) may lead to a poor fit of the detection function.

A further implicit assumption is that detection probabilities depend solely on the perpendicular distances of the objects from the transect line. If other factors, such as weather conditions or the size of the objects, affect detection probabilities, then standard line transect estimates of $f(0)$ may be biased. In addition, when size-bias is present, estimates of the expected cluster size based on the observed mean cluster size will also be biased.

Throughout this thesis it is assumed that assumption (ii) holds, so that standard line transect estimation procedures can be applied.

1.3 'Platforms of opportunity' and the use of design-based versus model-based methods of estimation

Line transect sampling is a design-based approach to abundance estimation, where the inclusion probabilities inherent to the sampling design form the basis of the estimator, thus avoiding the need for assumptions about the distribution of the population of interest (Hansen *et al.* 1983, Thompson 1992). To see this, it is useful to view the line transect estimators from expressions (1.6) and (1.7) from a Horvitz-Thompson (Horvitz and Thompson 1952) perspective. A Horvitz-Thompson estimator of the abundance of objects in the survey region, N, is given by (Thompson 1992):

$$\hat{N}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\rho_i}, \tag{1.10}$$

where y_i corresponds to the cluster size of the ith sampled object ($i = 1, \ldots, n$) and ρ_i denotes the probability that the ith object is included in the sample. For simplicity assume that objects correspond to single individuals, that is, $y_i = 1$, although the results below also apply to objects that occur in clusters. The estimator from the previous expression is then:

$$\hat{N}_{HT} = \sum_{i=1}^{n} \frac{1}{\rho_i}. \tag{1.11}$$

Note that $\hat{N}_{HT}$ is unbiased if the ρ_i are known.

In the context of line transect sampling, ρ_i is the probability that object i is within the area surveyed and is detected. From assumption (i) from the previous section, the probability that an object is in the area surveyed is given by $2LW/A$. Define p_i to be the probability that object i is detected, given that it is within the area surveyed, that is (Borchers 1996):

$$\begin{aligned} p_i &= \mathbb{E}\left[g(x_i)\right] \\ &= \int_0^W g(x)\,\pi(x)\,dx, \end{aligned} \tag{1.12}$$

where $g(x) = f(x)/f(0)$. Assuming that transect lines are randomly placed within the survey region, then $\pi(x) = 1/W$ and so:

$$\begin{aligned} p_i &= \frac{1}{W}\int_0^W g(x)\,dx \\ &= \frac{1}{Wf(0)}. \end{aligned} \tag{1.13}$$

Note that it is assumed that W is small relative to the size of the study area. If it is not, and W extends beyond it, the assumption that the objects are uniformly distributed relative to the transect line (*i.e.* $\pi(x) = 1/W$) is violated.

We then have

$$\begin{aligned} \rho_i &= \frac{2LW}{A} \cdot p_i \\ &= \frac{2L}{A \cdot f(0)}, \end{aligned} \tag{1.14}$$

which can be estimated by:

$$\hat{\rho}_i = \frac{2L}{A \cdot \hat{f}(0)}. \tag{1.15}$$

Substituting the above result into expression (1.11), a Horvitz-Thompson estimator of the abundance of objects in the survey region is given by:

$$\begin{aligned} \hat{N}_{HT} &= \frac{A}{2L} \cdot \sum_{i=1}^{n} \hat{f}(0) \\ &= A \cdot \frac{n \cdot \hat{f}(0)}{2L} \\ &= A \cdot \hat{D}, \end{aligned} \tag{1.16}$$

which corresponds to the standard line transect estimator from section (1.1). Note that, conditioning on n, the above estimator will be unbiased as long as $\hat{f}(0)$ is unbiased.

Although under the design-based approach prior knowledge about the target population is not directly incorporated into the estimator, it can be used to increase the efficiency of the estimator via the survey design (Hansen *et al.* 1983). For example, if it is known *a*

priori that the density of objects in the population varies according to the characteristics of the various habitats within the survey region, then this information may be used to stratify the survey region by habitat and optimise the allocation of effort across the various habitats to be surveyed (e.g. Buckland *et al.* 1993).

Dedicated design-based surveys, however, are often costly. A typical example is the case of surveys of marine mammals, which generally occur over vast areas that are expensive to survey. Thus opportunistic survey platforms, such as commercial vessels or ferries, are often used in place of dedicated survey vessels. Such platforms usually follow pre-determined routes that do not conform to any survey design, and which may not encompass areas with densities representative of the entire survey region. As a result, line transect estimates of abundance may be substantially biased. For example, if effort is concentrated in areas of high density, total population abundance will be overestimated. Nonetheless, data collected from opportunistic survey platforms often span relatively long periods of time, and can provide useful information on trends in abundance of the population of interest if the bias in the estimates can be eliminated. Even if bias cannot be completely eliminated, but instead minimised and maintained relatively constant throughout the time period of interest, trends in relative abundance may be used to assess the status of the population.

Bias in abundance estimates resulting from non-random search effort may be minimised through the use of either a design-based or a model-based approach. An example of the former is post-stratification (Anganuzzi and Buckland 1993), where strata are defined based on similarities between smoothed values of the variable of interest (or a proxy that correlates well with the variable of interest) over a grid of cells, and estimates of the value of the variable of interest are then obtained for each stratum. An overall estimate is obtained as an area-weighted average of the values from all strata, thus minimising the bias that results from a greater allocation of effort to some areas. The method can be applied to the many components of an estimator (e.g. for line transect sampling such components would be encounter rate, mean school size and effective strip (half-)width), with different stratifications used for each component. However, one limitation of this approach is that it may perform poorly if effort is very sparse or very concentrated within a small area (Anganuzzi and Buckland 1993).

In contrast, model-based procedures make use of knowledge about the distribution of the variable of interest and its relationship with auxiliary variables which are thought to affect it. Thus, even though data from opportunistic survey platforms may not cover

areas representative of the entire region, model-based estimators can recognise the trends in the response variable, thus minimising bias resulting from the non-random search effort. Values of the variable of interest where data are missing can then be predicted, allowing inference about the entire population to be made. Depending on the model used, the required sample sizes may be relatively small (Hansen *et al.* 1983). As with the post-stratification approach, different components of the estimator may be separately modelled. The above mentioned features rely on the assumption that the model is correct, in which case the estimator would be unbiased. However, if the model has been mis-specified, inference about the population being modelled may be misleading.

1.4 Thesis organisation

The remainder of this thesis is organised in four Chapters. Chapter 2 presents a model-based abundance estimation procedure that can be applied to line transect data, and which is particularly useful when survey effort is not random. Generalised additive models are used to separately model encounter rate and mean school size as a function of spatially and temporally referenced variables, and later used to generate a density surface for the entire region. An overall estimate of abundance is obtained by numerically integrating the density surface over the entire study area. The method relies on the assumption that $f(0)$ is constant throughout the region, as sample sizes usually preclude separate estimation of $f(0)$ across space and time. This limitation may be overcome through the application of the methodology proposed in Chapter 3, which presents a conditional likelihood approach that allows the effects of covariates on the detection function to be modelled. The estimated model parameters can then be used to obtain estimates of $f(0)$ which vary in space and time. Estimates of abundance can then be obtained either via standard line transect estimation or through the use of a Horvitz-Thomson-like estimator. Chapter 4 looks at the estimation of trends in abundance, including a description of an approach for assessing the status of populations which does not rely on knowledge about the existence and/or distribution of "stocks". Finally, Chapter 5 contains a general discussion of the work presented in the thesis and summarises the main results.

Motivation for the work presented here stemmed from the desire to develop alternative methods to obtain estimates of eastern tropical Pacific dolphin relative abundance based on sightings data collected by observers placed on board tuna vessels. These data are used in all the examples presented in the thesis, hence a detailed description is contained in Appendix A.

1.5 Appendices

1.5.1 Estimation of sample sizes required by the Petersen method

Seber (1982) formulates the precision of abundance estimates $\hat{N}$ based on the probability that the estimate will not differ from the true population size N by more than $100A\%$; that is:

$$1-\alpha \leq \Pr\left[-A < \frac{\hat{N}-N}{N} < A\right].$$

Choice of values for α and A will depend on the objectives of a given study, but Seber recommends $1-\alpha = 0.95$ and $A = 0.50$ for when only rough estimates of abundance are required, and $1-\alpha = 0.95$ and $A = 0.25$ for management purposes.

Given α and A, define:

$$1-\alpha = \phi\left(\frac{A\sqrt{D}}{1-A}\right) - \phi\left(\frac{-A\sqrt{D}}{1+A}\right), \tag{1.17}$$

where $\phi(\cdot)$ denotes the cumulative standard normal distribution, and let:

$$D = \frac{n_1 n_2 (N-1)}{(N-n_1)(N-n_2)},$$

where n_1 and n_2 denote the number of individuals sampled at each capture occasion. Solving expression (1.17) for D, we can then solve the expression above by assuming $n_1 = n_2 = n$.

Seber (1982, p.67) provides values of D for selected α and A. In the case of populations greater than 100 individuals, $D = 24.4$ for $\alpha = 0.05$ and $A = 0.50$, and $D = 69.9$ for $\alpha = 0.05$ and $A = 0.25$. Thus given a value for N of 500 000 individuals, we get $n = 3\ 469$ and $n = 5\ 843$ for each of the two scenarios mentioned, respectively.

1.5.2 Estimation of the expected number of detections required by the line transect estimator

In the context of line transect sampling, we use the coefficient of variation, $CV[\hat{N}]$, of the abundance estimate $\hat{N}$ as our measure of precision, where $CV[\hat{N}] = \widehat{SE}(\hat{N})/\hat{N}$. Given initial estimates of encounter rate (n_0/L_0), $f(0)$ and the expected cluster size of the population ($\mathbb{E}[s]$), and their respective variances (*i.e.* $\widehat{var}\,(n_0)$, $\widehat{var}\left\{\hat{f}(0)\right\}$ and $\widehat{var}(\hat{\mathbb{E}}[s])$), the total line length L required for a certain target level of precision ($CV_t[\hat{N}]$) to be achieved is given by (Buckland *et al.* 1993, p.303-306):

$$L = \frac{\hat{b} + \left\{\widehat{CV}[\hat{\mathbb{E}}[s]] \cdot \sqrt{n}\right\}^2}{\{CV_t[\hat{N}]\}^2} \cdot \frac{L_0}{n_0},$$

where $\hat{b} = \{\widehat{var}(n_0)/n_0 + n_0 \cdot \widehat{var}\{\hat{f}(0)\}/\{\hat{f}(0)\}^2\}$. Solving the above expression, the estimated number of detections, given L, can be obtained by solving $L_0/n_0 = L/n$ for n.

We use estimates of encounter rate, expected cluster size and effective strip half-width $(1/f(0))$ for eastern spinner dolphins in 1985 (Table 18 in Buckland and Anganuzzi 1988*b*), whose estimated population size $\hat{N}$ is 570 000 (although this is an estimate of *relative* rather than absolute abundance, the estimated parameters will serve for our illustrative purposes). Following the procedure described above, we have $\hat{b} = 13.0769$ and, given $n_0/L_0 = 2.4$, an estimated number of detections $n = 94$ and $n = 378$ for a $CV_t[\hat{N}]$ of 50% and 25%, respectively.

Chapter 2

Spatio-temporal modelling of line transect data from opportunistic surveys

2.1 Introduction

In the previous Chapter it was recognised that opportunistic survey platforms often provide the only means to obtain data on species that occur over vast areas, for which dedicated design-based surveys are too costly. However, application of line transect methodology to such data are problematic, as the non-random survey effort will likely introduce bias in the resulting line transect estimates of abundance. If effort is concentrated in areas where densities are high, for example, abundance estimates will be biased upwards. Anganuzzi and Buckland (1993) proposed a post-stratification procedure that minimises bias associated with non-random search effort (*cf.* Chapter 1), and applied the method to estimate encounter rate, mean school size and effective strip (half-)width using eastern tropical Pacific dolphin sightings data obtained by observers placed on board tuna vessels (Anganuzzi and Buckland 1989). However, their method may perform poorly in areas with little or no effort, as values of the variable of interest may have to be smoothed over large distances, within which the values of the variable of interest may vary considerably.

In this Chapter we present a model-based approach to abundance estimation based on the application of generalised additive models (Hastie and Tibshirani 1990) to line transect data. As an example we apply the method to estimate relative abundance of eastern tropical Pacific dolphins using the tuna vessel observer data previously analysed by Buckland and Anganuzzi (1988*b*) and Anganuzzi and Buckland (1989), and described in detail in

2.2 Spatial and spatio-temporal generalised additive models

Generalised additive models (GAMs) have been widely applied in ecological studies to describe the spatial relationship between some quantity of interest, such as density or abundance, and a set of explanatory variables (see Augustin (1999) for a review). Examples include the modelling of the effects of oceanographic and bathymetric variables on fish density and/or abundance (Swartzman *et al.* 1994, 1995, Welch *et al.* 1995, Swartzman 1997); the examination of vegetation type distributional patterns as a function of topographic and biophysical variables (Brown 1994); and the assessment of the influence of environmental variables on the distribution and abundance of cetaceans (Forney 1997, 1999).

A second application of spatial and spatio-temporal GAMs has been in abundance estimation, where the estimated parameters from the spatially and temporally referenced GAM can be used to predict a surface for the response variable over areas which have not been surveyed. The predicted surface may be the quantity of interest itself (e.g. density or abundance), or it may be some quantity which is required for abundance estimation. An example of the latter is given by Swartzman *et al.* (1992), who used GAMs to estimate the mean catch-per-unit-effort (CPUE) at each survey point for a number of groundfish species in the Bering Sea. The estimated CPUE values can be later rescaled and summed, to provide an estimate of fish abundance over the desired stratum. Another example is given by the work of Borchers *et al.* (1997*a*,*b*) and Augustin *et al.* (1998), who applied GAMs with spatially and temporally referenced covariates to model the abundance of fish eggs. Hedley *et al.* (1999) and Hedley (2000) developed spatio-temporal models for line transect data using a GAM framework. She obtained estimates of cetacean abundance by modelling either the number of detected animals within transect segments or the detected areas between sightings.

An underlying assumption of all of the above models is that any spatial correlation in the response variable has been explained by the covariates included in the models (Augustin *et al.* 1996), something which can be assessed through either formal (statistical tests) or informal (graphical examination) means using the models' residuals. However, this poses difficulties for the interpretation of the underlying relationship between the response variable and the covariates, as it is not always possible to separate the covariate effects from

spatial autocorrelation (Augustin 1999). For example, the presence of autocorrelation in the model residuals may reflect a poor model fit, rather than any real spatial dependence in the response. Conversely, the absence of any autocorrelation may simply indicate that such autocorrelation has been absorbed by one or more covariates in the model. One way of minimising this problem is to identify the spatial resolution for the response variable at which it does not show any spatial dependence on neighbouring values (e.g. Swartzman *et al.* 1992). In the case where the model is used mainly to predict a surface for the response, however, this problem is not as critical since the main objective is prediction, rather than inference. Under such circumstances, it suffices to ensure that the model residuals do not exhibit any spatial dependence, as the approach is robust to the presence of spatial dependence (Augustin 1999).

2.3 GAM framework

GAMs are an extension of generalised linear models (GLMs; McCullagh and Nelder 1989), in which the response variable is modelled as a smooth function of one or more explanatory variables (Hastie and Tibshirani 1990):

$$\mu_i = g^{-1}\left(\beta_0 + \sum_{k=1}^{q} f_k\left(u_{ki}\right)\right). \tag{2.1}$$

Here μ_i $(i = 1,\ldots,n)$ denotes the expectation of independent and identically distributed (iid) random variables y_i which follow one of the distributions from the exponential family. $\eta_i = \beta_0 + \sum_{k=1}^{q} f_k\left(u_{ki}\right)$ is an additive predictor which is a function of the explanatory variables u_{ki} $(k = 1,\ldots,q)$, with β_0 as the intercept and where $f_k(\cdot)$ denotes a smooth function of the kth covariate u_k. The μ_i are related to the additive predictor η_i via the link function $g(\cdot)$, so that $g(\mu_i) = \eta_i$ (Hastie and Tibshirani 1990). Choice of link function will depend on the specified distribution of the y_i and the desired structure for the model.

Parameters of the above model can be estimated by minimising the penalised residual sum of squares (PRSS):

$$\text{PRSS} = \sum_{i=1}^{n}\left\{y_i - \sum_{k=1}^{q} f_k(u_{ik})\right\}^2 + \sum_{k=1}^{q}\lambda_k \int \left\{f_k''(t)\right\}^2 dt. \tag{2.2}$$

The first term is akin to the residual sum of squares from standard regression, which gives

a measure of how well the model fits the data. The second term penalises the curvature in the function, with the constant λ_k denoting the amount of smoothing of the function $f_k(\cdot)$. Note that if λ_k is set to zero, the second term vanishes and the solution to the above equation results in a function $f_k(\cdot)$ which interpolates the data. If, on the other hand, λ_k approaches infinity, the penalty term will also approach infinity unless $f_k(t)$ is linear, so that $f_k''(t) = 0$ (Hastie and Tibshirani 1990). The advantage of GAMs over other regression-based approaches lies in this flexibility of the smoother $f(\cdot)$, which is not restricted to having a linear form.

A number of functional forms for $f(\cdot)$ may be specified, all based on the averaging of observations over some neighbourhood around a target value (see Hastie and Tibshirani (1990) for a comprehensive review). The way this averaging is done depends on the choice of smoother. In the simplest cases, straight averages are used. For example, given a set of observations (x_i, y_i) $(i = 1, \ldots, n)$, a smoother $f(x)$ may correspond to the average of the observations y_i that fall within neighbourhoods defined by the values of x_{j-1} and x_j $(j = 1, \ldots, m)$ such that $x_{j-1} < x_i \leq x_j$ (bin smoother), or to the average of k observations y_i to the left and to the right of each data point x_i (running mean). Kernel smoothers, on the other hand, compute a weighted mean of the y_i over some neighbourhood around the target value x_0 according to some specified function (kernel), in a way such that the weight applied to each y_i decreases the further away its corresponding value of x_i is from x_0. A common choice of kernel is the Gaussian density.

The remaining smoothers are all based on piecewise polynomials (splines) applied to regions defined by knots along the x-axis, and which are forced to have continuous first and second derivatives at the knots, so that they join smoothly at those knots. Both regression splines and cubic smoothing splines are based on piecewise cubic polynomials, but they differ in the way they are constructed. While the former requires the specification of both the number and location of the knots, cubic smoothing splines are directly obtained as the minimiser of the PRSS (Appendices 2.7.1 and 2.7.2).

Hastie and Tibshirani (1990) point out that there are few explicit comparisons of the various smoothers in the literature, and hence no single smoother may be recommended over the others. However, kernel smoothers and smoothing splines generally provide a smoother fit than other smoothers based on straight averaging. An exception is the locally-weighted running-line smoother (loess) of Cleveland (1979), which gives the smoothness of kernels and splines via the smoothing of weighted averages of observations over some specified neighbourhood. We chose to use cubic smoothing splines in all models presented

in this Chapter, with the degree of smoothing of each term specified by its effective number of parameters or 'degrees of freedom' (*df*; Hastie and Tibshirani 1990). The latter ranges from 1 *df* (a linear term) to a number of *df* that approaches the number of observations (and hence interpolates the data).

Model selection in GAMs involves both the choice of covariates to be included in the model and the degree of smoothing of each smooth covariate term. Clearly the number of potential models to be selected is prohibitively large, as the number of combinations of which covariates to include in the model and the *df* for each covariate is great. One way of narrowing the potential number of models to be considered is to constrain the number of *df* to be applied to each covariate. As an example, Fewster and Buckland (1996) fitted GAMs to annual estimates of bird abundance as a function of time (year), and selected a cut-off point for the *df* of the covariate term based on the wiggliness of the fit. Since the interest in that case was to model long term trends, rather than annual fluctuations in abundance, models with a small degree of smoothing (i.e. large *df*) that reproduced much of the noise in the annual estimates of abundance were ruled out. They then applied model selection criteria to choose a final model amongst the restricted set of models. This is the general approach we adopted.

2.4 Modelling line transect data using GAMs

Let there be a set of data obtained from standard line transect surveys conducted over some survey region of size A. If surveys were carried out according to some survey design, and all assumptions of line transect methodology held, then standard line transect estimation procedures could be applied to yield an estimate of the abundance of the objects of interest within the survey region A. Here we are concerned with the case where the survey was poorly designed, or data were collected from opportunistic platforms, so that survey effort was either not random, or coverage of the survey region was poor, or both. In such cases standard line transect estimation procedures would likely result in biased estimates of density, and hence of abundance, as estimates of encounter rate, the expected cluster size, and $f(0)$ (the inverse of the effective strip (half-)width) would not be representative of the entire survey region. In this Chapter we concentrate on the estimation of encounter rate and the expected cluster size, deferring the problem of estimating $f(0)$ to Chapter 3.

Given the poor coverage of the survey region, some areas will not have been surveyed, and so no data with which to estimate encounter rate and the expected cluster size are

available for those areas. Hence we need to model those quantities, so that an estimate of abundance over the entire survey region may be obtained. In addition, surveys may also exhibit poor coverage in time, in which case our models must be both spatially and temporally referenced. We use the GAM framework described in the previous section to model encounter rate and the expected cluster size, estimated by the observed mean cluster size, in space and time over the entire survey region as a function of a set of spatially and temporally referenced explanatory variables.

To make the model from expression (2.1) spatially and temporally explicit, define a spatio-temporal grid of cells indexed by lt, where l ($l = 1, \ldots, m$) specifies each cell in space and t ($t = 1, \ldots, T$) in time. For each cell we compute estimates of encounter rate (n_{lt}/L_{lt}) and the mean cluster size of detected objects (s_{lt}) using all observations within each cell. Here n_{lt} denotes the total number of detected clusters and L_{lt} is the total amount of effort surveyed. If we had observations for all $m \times T$ cells, we would not have to model the data. Hence assume that not all $m \times T$ cells contain observations, although values for the explanatory variables $\mathbf{u}_{lt}$ ($\mathbf{u}_{lt} = u_{1lt}, \ldots, u_{qlt}$) must be available for all cells.

2.4.1 Modelling encounter rate

In order to apply GAMs to model encounter rate, we need to specify a distribution for that quantity. However, it is not clear what distributional form encounter rate may exhibit. Instead we model the number of detected clusters n_{lt} within each cell according to a Poisson distribution and using a log link function. Since the encounter rate for cell lt is defined as n_{lt}/L_{lt}, where L_{lt} denotes the total search effort expended within that cell, we can then include the logarithm of the latter quantity as an 'offset' in the model, so that it is effectively treated as a constant. Hence we have:

$$\mathbb{E}[n_{lt}] = \exp\left\{\log_e(L_{lt}) + \beta_0 + \sum_{k=1}^{q} f_k(u_{klt})\right\}, \tag{2.3}$$

which can be written as:

$$\log_e\left\{\frac{\mathbb{E}[n_{lt}]}{L_{lt}}\right\} = \beta_0 + \sum_{k=1}^{q} f_k(u_{klt}). \tag{2.4}$$

Note that if the survey period is short enough so that encounter rates are unlikely to

have changed in space over time, the latter can be omitted from the model so that data across all time periods can be pooled. Otherwise, one of the explanatory variables in the model must be some measure of time, such as month or year, so as to allow estimates of abundance at different times t, and also overall (averaged over the entire survey period) to be obtained (see section 2.4.3).

The Poisson distribution assumes that the variance is equal to the mean. However, if the number of detected clusters within the cells shows great variability, then the variance will be greater than the mean; this is commonly referred to as overdispersion. In such cases, the use of the negative binomial distribution may be more appropriate (*cf.* Venables and Ripley 1994).

2.4.2 Modelling mean cluster size

The mean cluster size within a given cell is a continuous quantity, although its underlying distribution is not clear. Hence we follow the approach described for encounter rate, so that our response variable is the total number of objects detected at location l and time t ($N_{lt} = \sum_{lti=1}^{n_{lt}} s_{lti}$), with the offset being the logarithm of the total number of detected clusters within that cell (n_{lt}). Assuming a Poisson distribution with a log link function, we then have:

$$\mathbb{E}[N_{lt}] = \exp\left\{\log_e(n_{lt}) + \beta_0 + \sum_{k=1}^{q} f_k(u_{klt})\right\}, \tag{2.5}$$

which can be expressed as:

$$\log_e\left\{\frac{\mathbb{E}[N_{lt}]}{n_{lt}}\right\} = \beta_0 + \sum_{k=1}^{q} f_k(u_{klt}) \tag{2.6}$$

and where N_{lt}/n_{lt} corresponds to the mean cluster size s_{lt} in cell lt.

Like the modelling of encounter rate, in the presence of overdispersion the negative binomial distribution may be used in place of the Poisson.

2.4.3 Density and abundance estimation

Given the estimated parameters from the encounter rate and mean cluster size GAMs, we can then obtain a predicted surface for each of those quantities for all $m \times T$ cells. Assuming an estimate of $f(0)$ is available, an estimate of the density of objects at spatial location l and time t ($\hat{D}_{lt}$) is obtained using standard line transect methodology (Buckland *et al.* 1993):

$$\hat{D}_{lt} = \frac{1}{2} \cdot \hat{f}(0) \cdot \frac{\widehat{n_{lt}}}{L_{lt}} \cdot \hat{s}_{lt}. \tag{2.7}$$

An estimate of the overall abundance of objects at time t ($\hat{N}_t$) can then be obtained by numerically integrating the predicted density surface at time t over the entire survey region. As it is implicitly assumed that the predicted density estimate for each cell is constant within each cell, this effectively corresponds to the sum of the abundance estimates from all m cells at time t; that is:

$$\begin{aligned} \hat{N}_t &= \sum_{l=1}^{m} \hat{N}_{lt} \\ &= \textstyle\sum_{l=1}^{m} A_l \cdot \hat{D}_{lt} \\ &= \frac{\hat{f}(0)}{2} \cdot \sum_{l=1}^{m} A_l \cdot \hat{P}_{lt}, \end{aligned} \tag{2.8}$$

where A_l denotes the size of the lth cell and $\hat{P}_{lt} = (n_{lt}/L_{lt}) \cdot \hat{s}_{lt}$.

If the survey period extends over, say, 12 months, an overall estimate of abundance may be obtained as the average of the monthly estimates of abundance:

$$\hat{N} = \frac{\sum_{t=1}^{T} \hat{N}_t}{T}, \tag{2.9}$$

where in this case $T = 12$.

2.4.4 Variance estimation

Estimates of the variance for the encounter rate and mean school size predicted surfaces from the GAMs cannot be directly obtained from the fitted models, as it is not possible to obtain estimates of the standard error for cells which did not originally have any observations. In addition, analytic estimates of the variance would rely on the assumption that observations from different cells were independent, something which is clearly not valid due to spatial correlation between values from neighbouring cells.

Instead, a bootstrap procedure (Efron and Tibshirani 1993) can be used to obtain an estimate of the variance of the final estimate of abundance, as follows. At each of B bootstrap iterations transect lines are resampled, with replacement, and the modelling of encounter rate and mean school size carried out as described in sections 2.4.1 and 2.4.2. We use transect lines as our resampling units so that observations from different transect lines can be assumed to be independent; if the observations themselves were resampled the assumption of independence might not hold. However, in opportunistic surveys where a survey design is not implemented, transect lines are often difficult to define. In such cases, it is important to choose the resampling units in a way such that observations from different units are independent, even though observations within a given unit may not be. To obtain an estimate of the abundance of objects, either one of two approaches can be adopted. If an estimate of $f(0)$ and its standard error is already available, then that estimate can be used to obtain an estimate of the abundance of objects at time t, $\hat{N}_t$, or, if desired, an overall estimate of abundance, $\hat{N}$, as described in section 2.4.3. An estimate of the standard error of the estimate of abundance can then be computed based on the B bootstrap estimates of $\hat{N}_t$ or $\hat{N}$, and an overall estimated coefficient of variation (CV) may be obtained as the square root of the sum of the squared CVs of the estimates of abundance and $f(0)$. Alternatively, at each bootstrap iteration an estimate of $f(0)$ based on standard estimation procedures (*cf.* Buckland *et al.* 1993) may also be computed, and an estimate of abundance obtained as previously described. Estimates of standard error can then be computed based on the B bootstrap estimates of $\hat{N}_t$ or $\hat{N}$. In addition, 'percentile' confidence intervals (Efron and Tibshirani 1993) can also be obtained from those bootstrap estimates.

2.5 Example: Spatio-temporal modelling of eastern tropical Pacific dolphin relative abundance

A number of authors have demonstrated strong correlations between certain oceanographic variables and dolphin distribution (Reilly 1990, Fiedler and Reilly 1994, Reilly and Fiedler 1994). We exploit these relationships to model encounter rate and mean school size of eastern tropical Pacific (ETP) dolphins as a function of spatially and temporally referenced oceanographic and locational variables. We use dolphin sightings data collected by observers placed on board tuna vessels (Appendix A). The method is applied to stocks of offshore spotted dolphins (*Stenella attenuata*), with stock boundaries located as described in Dizon *et al.* (1994). As we are interested in applying the results from our models to obtain estimates of dolphin relative abundance, we use annual estimates of $f(0)$ obtained using the approach described by Anganuzzi and Buckland (1989) (and available from Anganuzzi and Buckland 1994, IATTC 1994, 1995, 1996, 1997, 1998, 1999), so that $\hat{f}(0)$ is effectively constant in space within each year.

Dolphin sightings data for the ETP are available for the period 1975–1997. Potential oceanographic variables to be included in the models are available for 1° × 1° latitude–longitude cells for each month within 1975–1997. Hence we compute the total number of detected dolphin schools (n), the total number of dolphins sighted (N) and the total amount of effort surveyed (L) for each 1° × 1° latitude–longitude cell for each month. We followed the same data selection criteria used by Buckland and Anganuzzi (1988*b*), and applied a truncation distance of 5nm.

Due to the large size of the dataset, and potential difficulties in fitting a model to such a long time series, we modelled the data separately for each year. Potential oceanographic variables to be included in the models were sea surface temperature (*sst*), thermocline depth (*tcd*), thermocline strength (*tcs*) and a habitat suitability index (*hab*) which summarises the relationship between dolphin distribution and various oceanographic variables (Reilly and Fiedler 1994). Locational covariates included latitude (*lat*) and longitude (*lon*); and month (*mo*) was entered as a temporal index.

Thus our initial model for encounter rate for any given year was:

$$n_{lt} = \exp\{\log_e(L_{lt}) + \beta_0 + f(mo_{lt}; df) + f(lat_{lt}; df) + f(lon_{lt}; df) + f(sst_{lt}; df) + f(tcd_{lt}; df) + f(tcs_{lt}; df) + f(hab_{lt}; df)\} + \epsilon_{lt}, \tag{2.10}$$

where $f(\cdot; df)$ denotes a cubic smoothing spline with df degrees of freedom, $l = 1, \ldots, m$ corresponds to the spatial locations, and $t = 1, \ldots, 12$ corresponds to the months, and where ϵ_{lt} is the error term. To model mean school size, we also considered the inclusion of effort (eff) as a potential covariate in the model, since in areas of higher search effort schools may break up into smaller units. Our initial model for mean school size was then given by:

$$N_{lt} = \exp\{\log_e(n_{lt}) + \beta_0 + f(mo_{lt}; df) + f(lat_{lt}; df) + f(lon_{lt}; df) + f(sst_{lt}; df) + f(tcd_{lt}; df) + f(tcs_{lt}; df) + f(hab_{lt}; df) + f(eff_{lt}; df)\} + \epsilon_{lt}. \tag{2.11}$$

To restrict the potential number of df to be considered for each covariate, we started by fitting models which included all potential covariates with the same level of smoothing (*i.e.* same df). Possible df considered included 1 (*i.e.* linear terms), 2, 4, 6, 8 and 10. Examination of plots of the smoothed covariates showed that models with 4 and 6, and 8 and 10 df resulted in similar smooth functions. To minimise the possibility of spurious patterns in the data being reflected in the models, we chose to consider only the lowest degree of smoothing of each pair (*i.e.* 4 and 8 df). Given the large number of data points for each year, which includes much noise, we chose to include terms with 2 df in place of linear terms, as the latter appeared not to exhibit enough flexibility. Given the above set of potential df for each covariate, stepwise backward and forward selection was then carried out by varying the df of each smoother, beginning with a full model containing all possible covariates with an intermediate degree of smoothing (*i.e.* 4 df), and allowing for terms to be omitted from the model. The inclusion of any given covariate with a given df was determined based on Akaike's Information Criterion (AIC; Akaike 1973, Appendix 2.7.3).

Examination of diagnostic plots for the encounter rate models applied to offshore spotted dolphin sightings data revealed strong departures from normality in the residuals, suggesting that the Poisson distribution was not adequately modelling all the variability in

the data. A large proportion of the observed encounter rates had a value of zero, but a few observations exhibited very high values. As a result, the Poisson model tended to overfit the zeros whilst trying to fit those large encounter rate values, thus resulting in poor model fits. Some of the large encounter rate values observed were obtained in cells with very little effort. Based on the examination of the distribution of encounter rates from all cells, we discarded observations from cells in which encounter rates were greater than 0.5 schools sighted per nm searched. This resulted in a maximum of 0.34% of the observations from any given year being discarded (ranging from 0.03% to 0.34%, with mean 0.14%). Nonetheless, the Poisson model still did not adequately fit the data. The use of the negative binomial distribution (discussed in the next paragraph) in place of the Poisson yielded similar results. Hence we chose to follow a two-stage modelling approach, as follows. The presence or absence of any dolphin schools within each cell was modelled as a binomial distribution, with the probability of a school being present in any given cell being obtained from the fitted model. In addition, 13 "structural zeros" (Figure 2.1) were added to data from each month and each year, at the same locations; these are additional observations placed at locations just outside the stock boundaries, with the intent of preventing the model from extrapolating beyond the range of the data. Those structural zeros were placed in areas of suitable dolphin habitat, in which the models were consistently predicting unrealistically high values, but where the number of dolphin sightings was usually small (e.g. Dizon *et al.* 1994). We assigned 1000nm of effort and a value of zero sightings to each of those 13 cells; by assigning a relatively small amount of effort to these observations, their weight in the model fit is minimised, though they are still effective in preventing the model from predicting unrealistically large values near the outer boundaries of the survey region. However, the use of up to 8 *df* for the smooths in the models appeared to be contributing to the prediction of unrealistically high values near the outer stock boundaries. Therefore we restricted the *df* to 2, and used AIC to select whether to include a smoothed covariate term with 2 *df* or not. The total number of schools within each cell, conditional on at least one school having been recorded, was separately modelled by assuming a Poisson distribution and using observations from cells in which at least one school was observed. As with the binomial model, we used 13 structural zeros, except that in this case we assigned a value of one to the number of dolphin schools sighted in those cells. In addition, we also restricted the *df* to 2. However, for 1977, 1978 and 1996 we still obtained very poor model fits. In those years there were a few observations with very little effort but which resulted in extremely high encounter rates. Therefore, for those years we omitted those observations (which resulted in 0.24%, 0.12% and 0.05% of the observations deleted for each year, respectively) and did not use

any structural zeros, and let the *df* of the smooth covariate terms vary between 2, 4, or 8. A predicted encounter rate surface was finally obtained by multiplying the predicted surfaces from the models for presence/absence and for the total number of schools.

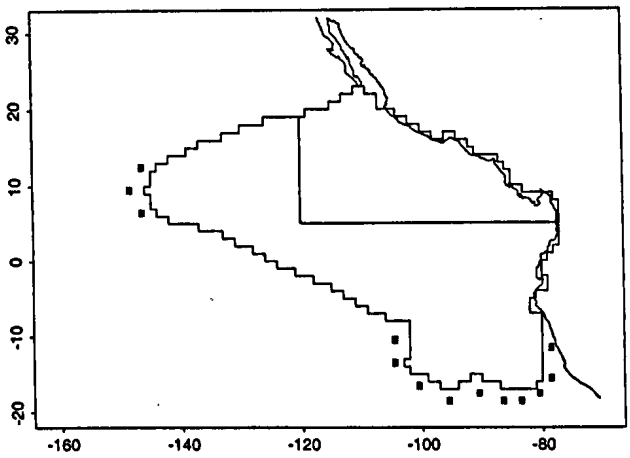


Figure 2.1: Location of structural zeros. Solid lines indicate stock boundaries for offshore spotted dolphins. See Section 2.5 for details.

Initial fitting of the mean school size model using a Poisson distribution resulted in extremely large (> 100) estimates of the dispersion parameter, which is assumed to be 1 for this distribution (so that the variance is equal to the mean). Examination of diagnostic plots for these models showed that they failed to provide good model fits for observations near the upper extreme of the range of the distribution. Thus we used instead the negative binomial distribution, which allows for extra variability via the scale parameter ϕ, where $var(Y) = \mu + \mu^2/\phi$. Maximum likelihood estimation of ϕ in the context of GAMs, however, is not feasible due to the iterative nature of the backfitting algorithm used to obtain the estimates of the smooth functions $f(\cdot)$. Hence we started by fitting a GLM with a negative binomial distribution to the data, including all potential covariates, and obtaining an estimate of ϕ. We then fitted a GAM with a negative binomial distribution, conditioning on our estimate of ϕ obtained from the initial GLM model. Model selection was then carried out as previously described, with $\hat{\phi}$ held constant. A predicted surface for mean school size was then obtained based on the final selected model.

Density estimates were obtained as described in section 2.4.3, and estimates of relative abundance for the northeastern and southern-western stocks of offshore spotted dolphins, as well as for both stocks combined, were obtained by integrating the density surface within the corresponding stock boundaries. Estimates of the variance of the annual estimates of abundance were obtained by bootstrapping. However, in the case of the tuna

vessel observer data, transects are difficult to define. Hence we chose individual cruises as our resampling units, which can assumed to be independent, and followed the balanced bootstrap (*cf.* Davison and Hinkley 1997) approach adopted by Buckland and Anganuzzi (1988*b*), in which each cruise appears the same number of times over all bootstrap iterations, although it may appear more than once, or not at all, in any given iteration. At each bootstrap iteration, encounter rate and mean school size were modelled as described above, conditioning on the final model originally selected. In the case of mean school size, in which a negative binomial GAM was used, an estimate of the scale parameter ϕ was first obtained by using a GLM. Estimates of dolphin density and abundance were obtained as previously described, using estimates of $f(0)$ from the literature, and bootstrap estimates of standard error were then computed based on the bootstrap estimates of abundance. Log-normal confidence intervals for each of the annual estimates of dolphin relative abundance were then obtained. Due to the large computational cost, a total of 79 bootstrap iterations were carried out for each year.

Note that in the bootstrap procedure described above we did not re-estimate $f(0)$, as estimates of its standard error are already available (Buckland and Anganuzzi 1994, IATTC 1994, 1995, 1996, 1997, 1998, 1999). Hence our bootstrap estimates of standard error were for $\left\{\hat{N}/\hat{f}(0)\right\} \cdot \hat{f}(0)$, and not for $\hat{N}$. To obtain an estimate of the standard error for $\hat{N}$, define $\hat{N} = \hat{N}_f \cdot f(\hat{0})$, where $\hat{N}_f = \hat{N}/\hat{f}(0)$. We can then use the delta method (*cf.* Seber 1982) to obtain an estimate of the CV of $\hat{N}$:

$$\left\{CV\left[\hat{N}\right]\right\}^2 = \left\{CV\left[\hat{N}_f\right]\right\}^2 + \left\{CV\left[\hat{f}(0)\right]\right\}^2. \tag{2.12}$$

From the bootstrap we can compute $\widehat{var}_b\left\{\hat{N}_f\right\}$; therefore, an estimate of $\left\{CV\left[\hat{N}_f\right]\right\}^2$ can be obtained as:

$$\left\{CV\left[\hat{N}_f\right]\right\}^2 = \frac{\widehat{var}_b\left\{\hat{N}_f\right\}}{\hat{N}_f^2}, \tag{2.13}$$

from where we have:

$$var(\hat{N}) = \hat{N}^2 \left\{ \frac{\widehat{var}_b\left\{\hat{N}_f\right\}}{\left\{\hat{N}_f\right\}^2} + \frac{\widehat{var}\left\{\hat{f}(0)\right\}}{\left\{\hat{f}(0)\right\}^2} \right\}. \tag{2.14}$$

To examine trends in relative abundance of both offshore spotted dolphin stocks, we applied a GAM with a Gamma distribution and a log link function to the annual point estimates of dolphin relative abundance, using year as a covariate. We used AIC to select the amount of smoothing in the model, letting the *df* vary from 1, 2, 4 or 8. 85% log-normal confidence intervals were computed based on the standard errors of the GAM model. Confidence intervals for the most recent year were then used to test for significant differences (at the 5% level) in relative abundance over time (Buckland *et al.* 1992), based on whether any of the 85% confidence intervals from previous years overlapped or not with those from the most recent one.

The covariates included in the final models for each year, as well as their degree of smoothing, are shown in Tables 2.1, 2.2 and 2.3. For illustrative purposes, plots of the smoothed covariates as a function of the covariate values for the binomial, count and mean school size models for 1995 are shown in Figures 2.2, 2.3 and 2.4, respectively. Predicted surfaces for encounter rate, mean school size and density for the years 1975-1997 are shown in Figures 2.5, 2.6 and 2.7, respectively.

Estimates of relative abundance for all stocks of offshore spotted dolphin are shown in Table 2.4. Clearly the estimates for all stocks for 1977 are unrealistically large. These poor estimates were caused by a bad fit for the encounter rate models; however, attempts to fit other models to the data did not improve the results. This is likely a result of the limited coverage and data available for that year, and/or data quality issues associated with data from the early years (Lennert-Cody *et al.* In prep). Bootstrap estimates of standard error for all stocks were proportionally larger for the 1970s, decreasing with time. This is likely a result of the greater amount of data available for the later years, and also better coverage, which resulted in more precise estimates of abundance.

Trends in offshore spotted dolphin relative abundance over the years 1975-1997 are shown in Figures 2.8 and 2.9 for the northeastern and southern-western stocks, respectively. Due to their extremely large values, estimates for all stocks for 1977 were omitted from all plots. For comparison we also show estimates of relative abundance obtained using the post-stratification approach.

Year	*mo*	*lat*	*lon*	*sst*	*tcd*	*tcs*	*hab*
1975	4	8	8	4	–	8	4
1976	2	8	8	4	8	8	2
1977	4	8	8	8	8	4	8
1978	8	8	8	8	2	8	8
1979	2	8	8	–	8	8	8
1980	2	8	8	–	8	8	4
1981	8	8	8	8	8	8	–
1982	8	8	8	4	4	8	8
1983	8	8	8	–	8	–	8
1984	8	8	8	–	8	8	–
1985	8	8	8	8	8	8	4
1986	8	8	8	8	8	2	2
1987	8	8	8	8	8	4	8
1988	4	8	8	4	8	8	2
1989	8	8	8	8	8	4	4
1990	8	8	8	4	2	8	2
1991	8	8	8	8	8	–	–
1992	8	8	8	8	8	8	4
1993	4	8	8	4	8	8	2
1994	2	8	8	8	4	8	4
1995	8	8	8	8	8	8	8
1996	8	8	8	8	8	8	8
1997	8	8	8	–	8	4	8

Table 2.1: Covariates selected for the presence/absence models for offshore spotted dolphins for the years 1975 – 1997. Numbers indicate the degrees of freedom of the smooth term. Potential covariates included month (*mo*), latitude (*lat*), longitude (*lon*), sea surface temperature (*sst*), thermocline depth (*tcd*), thermocline strength (*tcs*), and a habitat suitability index (*hab*).

Year	*mo*	*lat*	*lon*	*sst*	*tcd*	*tcs*	*hab*
1975	2	2	2	–	2	2	2
1976	2	2	2	–	2	2	2
1977	4	8	8	8	8	4	8
1978	8	2	4	2	4	2	–
1979	2	2	2	–	–	2	2
1980	2	2	2	2	2	2	2
1981	2	2	2	–	2	2	2
1982	2	2	2	–	–	2	2
1983	2	2	2	–	–	2	2
1984	2	2	2	2	–	2	2
1985	2	2	2	2	–	2	2
1986	2	2	2	–	–	2	–
1987	2	2	2	2	2	2	–
1988	2	2	2	2	–	2	2
1989	2	2	2	2	2	2	2
1990	2	2	2	2	2	2	2
1991	2	2	2	2	2	2	2
1992	2	2	2	–	2	2	2
1993	2	2	2	–	2	–	2
1994	2	2	2	2	2	2	–
1995	2	2	2	2	2	2	2
1996	8	8	8	8	–	8	4
1997	2	2	2	2	2	2	2

Table 2.2: Covariates selected for the models based on the total number of offshore spotted dolphin schools sighted within each cell, conditional on at least one school having been recorded, for the years 1975 – 1997. Numbers indicate the degrees of freedom of the smooth term. Potential covariates included month (*mo*), latitude (*lat*), longitude (*lon*), sea surface temperature (*sst*), thermocline depth (*tcd*), thermocline strength (*tcs*), and a habitat suitability index (*hab*).

Year	*mo*	*lat*	*lon*	*sst*	*tcd*	*tcs*	*hab*
1975	2	–	–	–	–	4	2
1976	4	2	–	2	4	2	–
1977	2	8	2	8	–	–	4
1978	4	4	4	–	–	–	–
1979	4	4	–	4	–	4	–
1980	4	8	8	8	–	8	2
1981	2	8	8	–	4	–	–
1982	4	2	4	–	–	–	2
1983	8	2	–	4	–	–	–
1984	2	8	4	–	4	–	8
1985	2	4	4	–	–	–	2
1986	2	8	4	2	4	–	–
1987	4	8	8	2	8	4	8
1988	4	8	8	–	8	–	2
1989	2	2	8	8	4	8	4
1990	4	8	8	2	2	–	–
1991	4	4	4	2	8	–	–
1992	2	8	8	8	8	8	2
1993	4	8	8	4	–	2	2
1994	8	8	8	4	8	–	8
1995	8	8	8	8	–	4	8
1996	2	8	–	8	–	8	–
1997	8	8	8	–	2	2	8

Table 2.3: Covariates selected for the mean school size models for offshore spotted dolphins for the years 1975 – 1997. Numbers indicate the degrees of freedom of the smooth term. Potential covariates included month (*mo*), latitude (*lat*), longitude (*lon*), sea surface temperature (*sst*), thermocline depth (*tcd*), thermocline strength (*tcs*), and a habitat suitability index (*hab*).

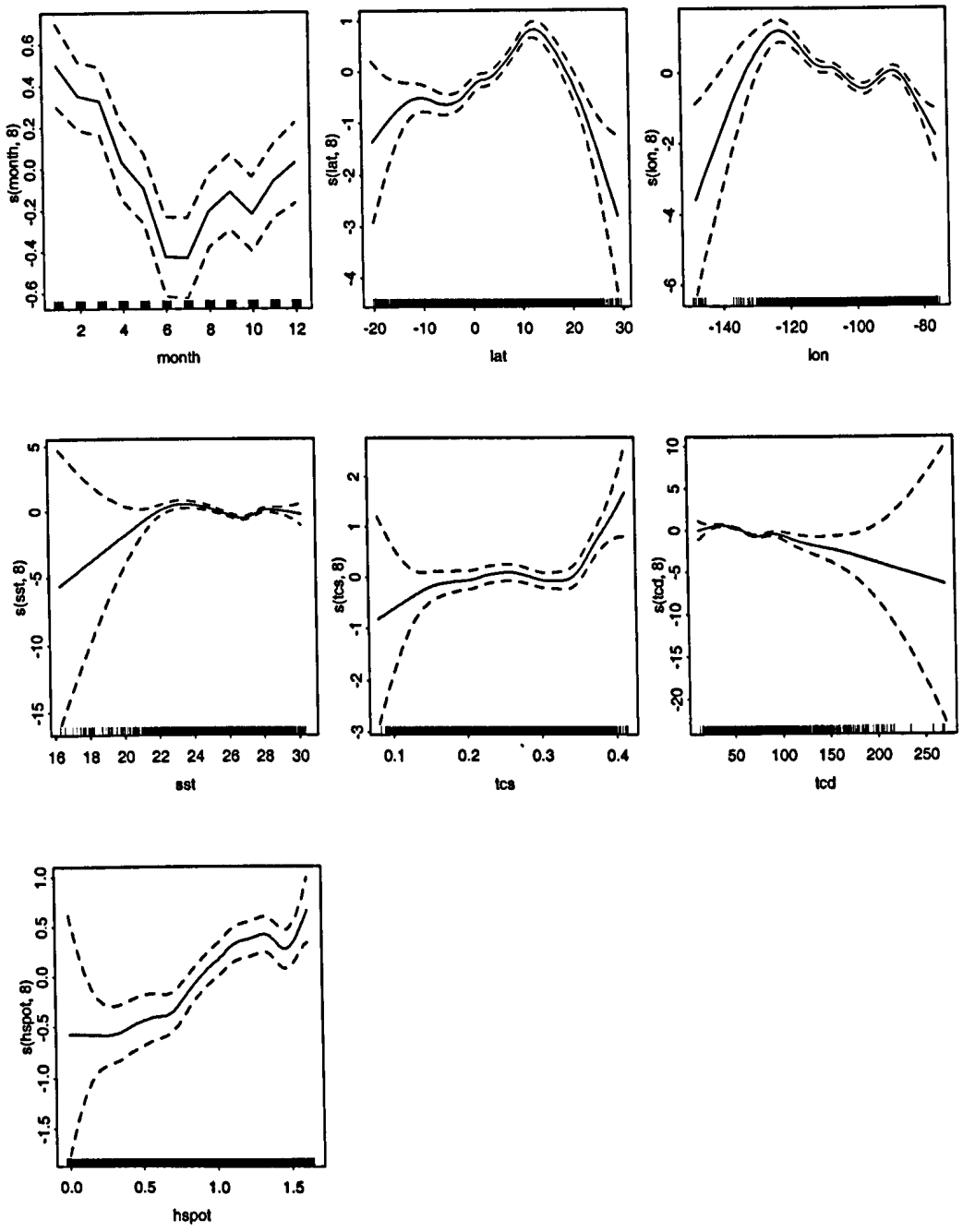


Figure 2.2: Plot of smoothed covariates, scaled relative to their mean value, against covariate values for the binomial model for the presence/absence of dolphin schools, for 1995. Solid lines indicate the smoothed covariate values; dashed lines correspond to two standard errors. Tick marks on the x-axes indicate observations. *mo* = month; *lat* = latitude; *lon* = longitude; *sst* = sea surface temperature; *tcd* = thermocline depth; *tcs* = thermocline strength; *hspot* = habitat suitability index.

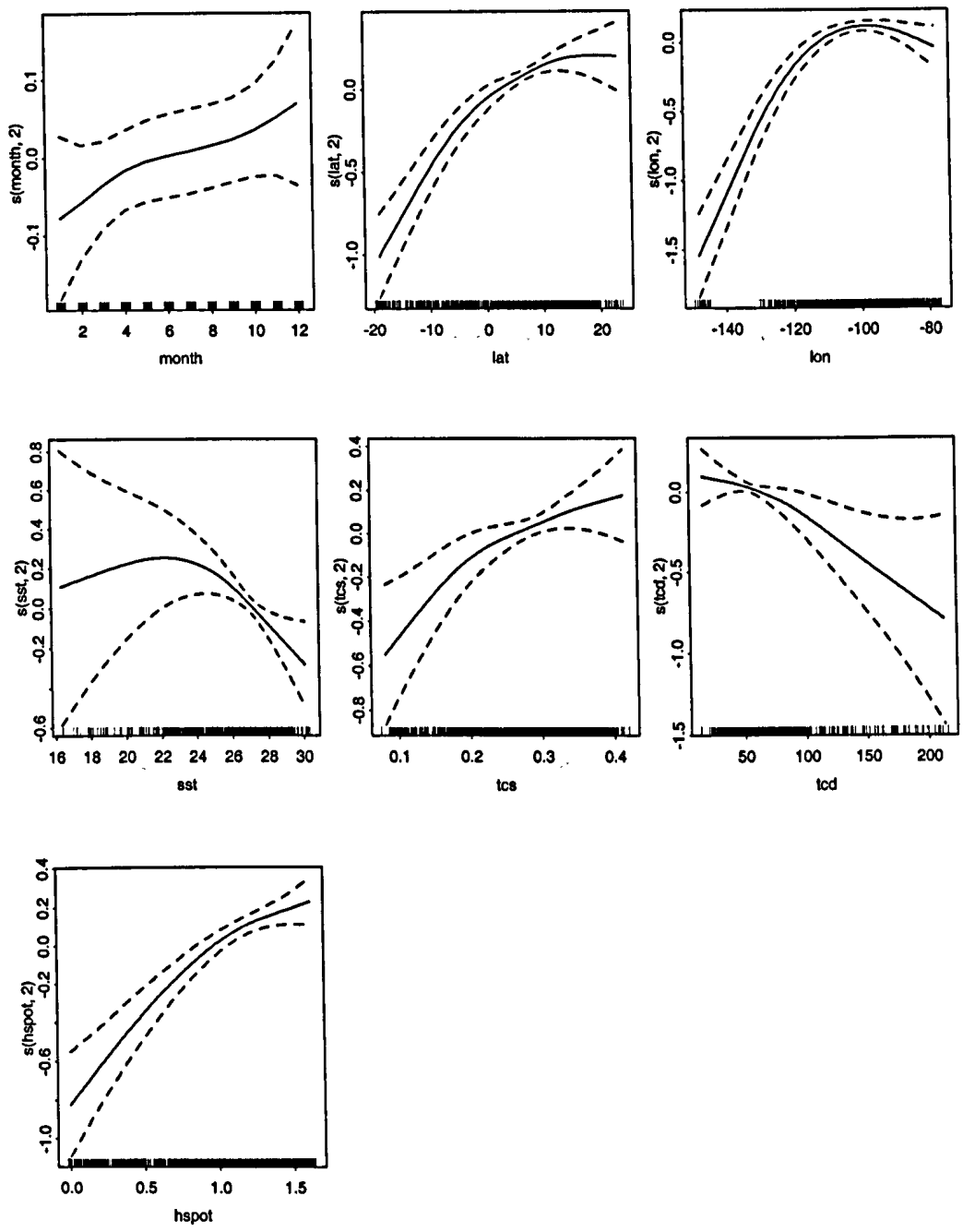


Figure 2.3: Plot of smoothed covariates, scaled relative to their mean value, against covariate values for the model for the number of dolphin schools, for 1995. Solid lines indicate the smoothed covariate values; dashed lines correspond to two standard errors. Tick marks on the x-axes indicate observations. *mo* = month; *lat* = latitude; *lon* = longitude; *sst* = sea surface temperature; *tcs* = thermocline strength; *tcd* = thermocline depth; *hspot* = habitat suitability index.

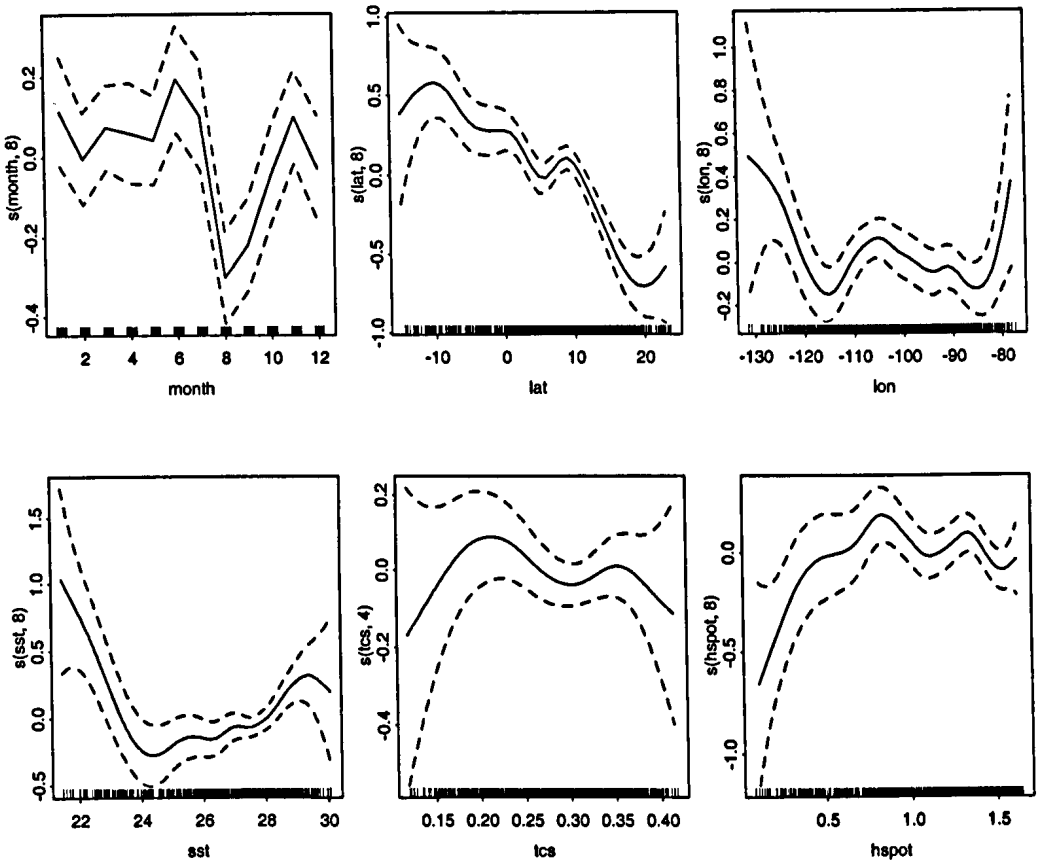


Figure 2.4: Plot of smoothed covariates, scaled relative to their mean value, against covariate values for the mean school size model for 1995. Solid lines indicate the smoothed covariate values; dashed lines correspond to two standard errors. Tick marks on the x-axes indicate observations. *mo* = month; *lat* = latitude; *lon* = longitude; *sst* = sea surface temperature; *tcs* = thermocline strength; *hspot* = habitat suitability index.

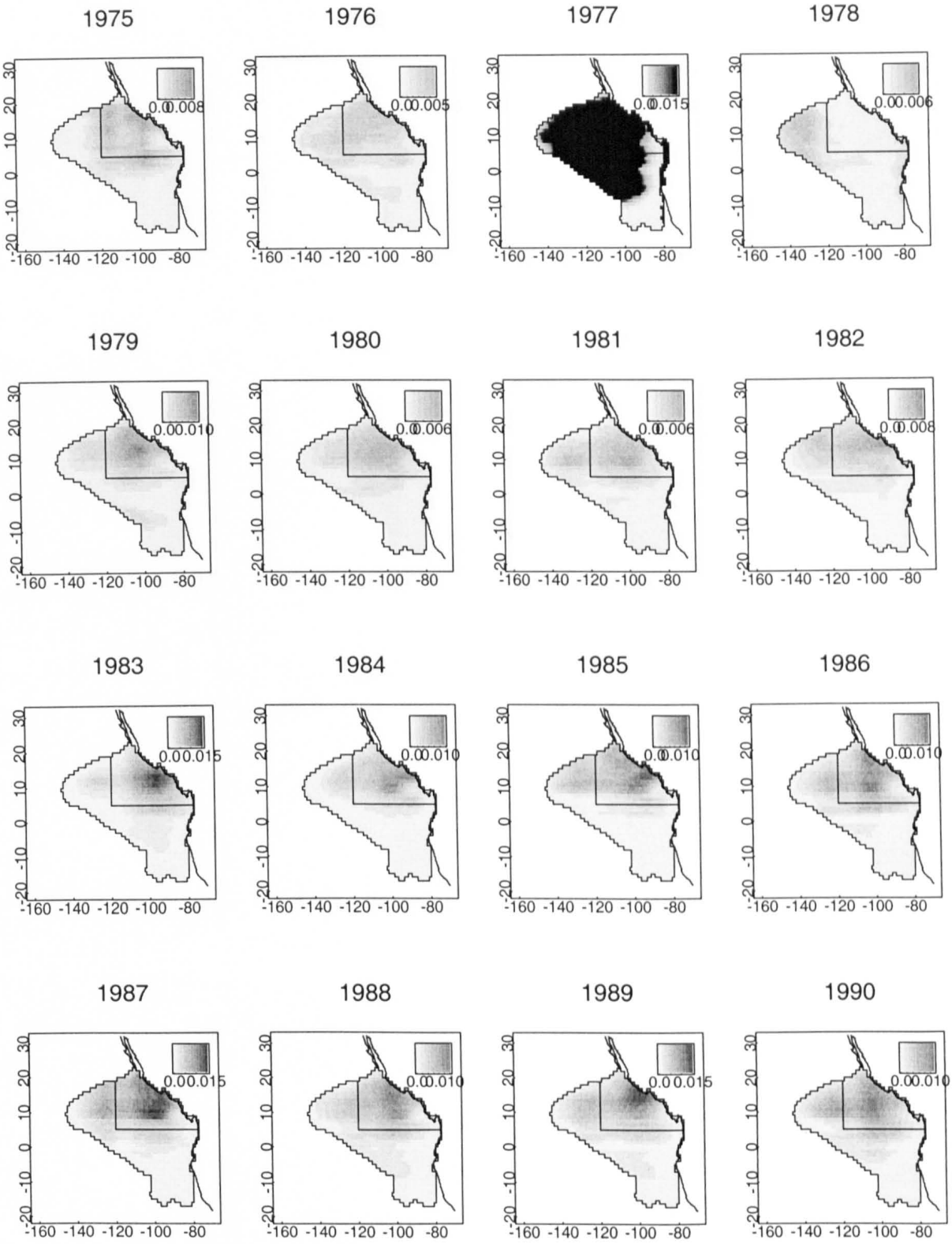


Figure 2.5: Predicted encounter rate surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.

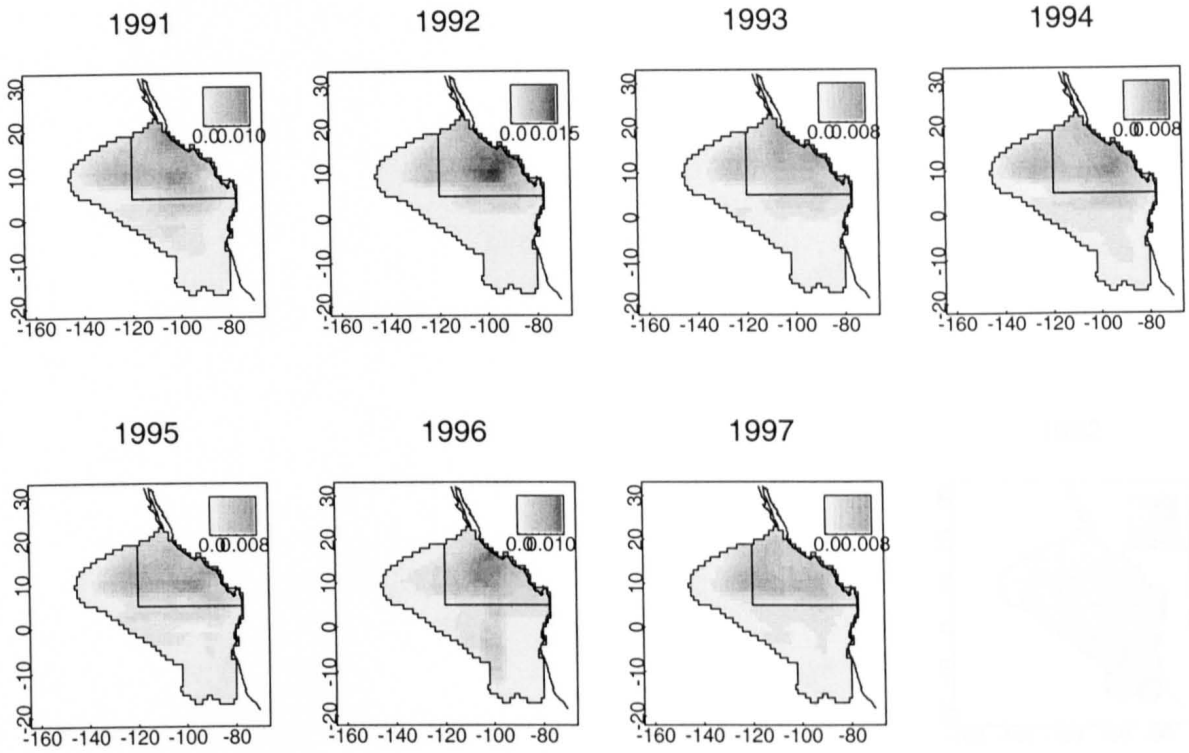


Figure 2.5: Predicted encounter rate surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.[continued]

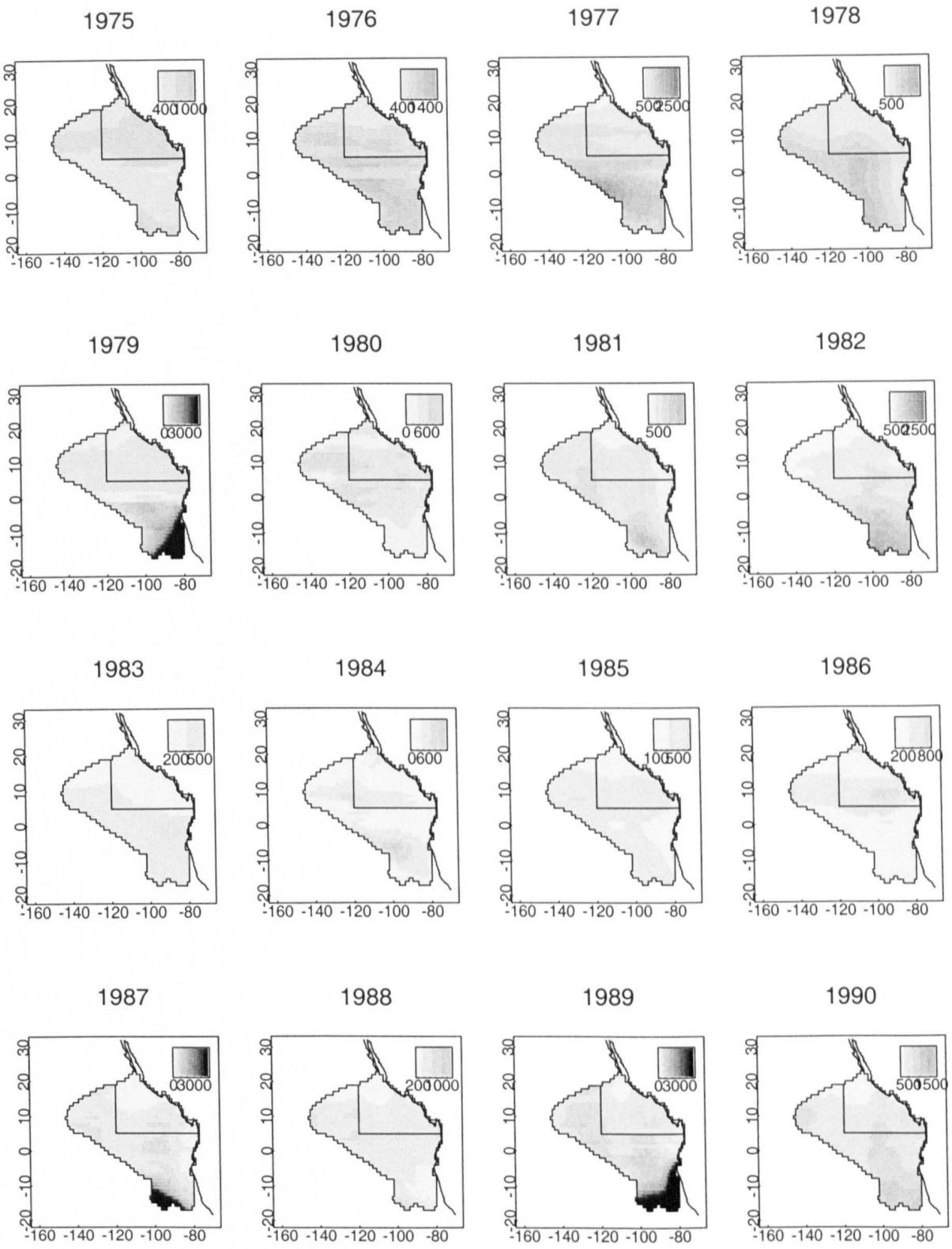


Figure 2.6: Predicted mean school size surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.

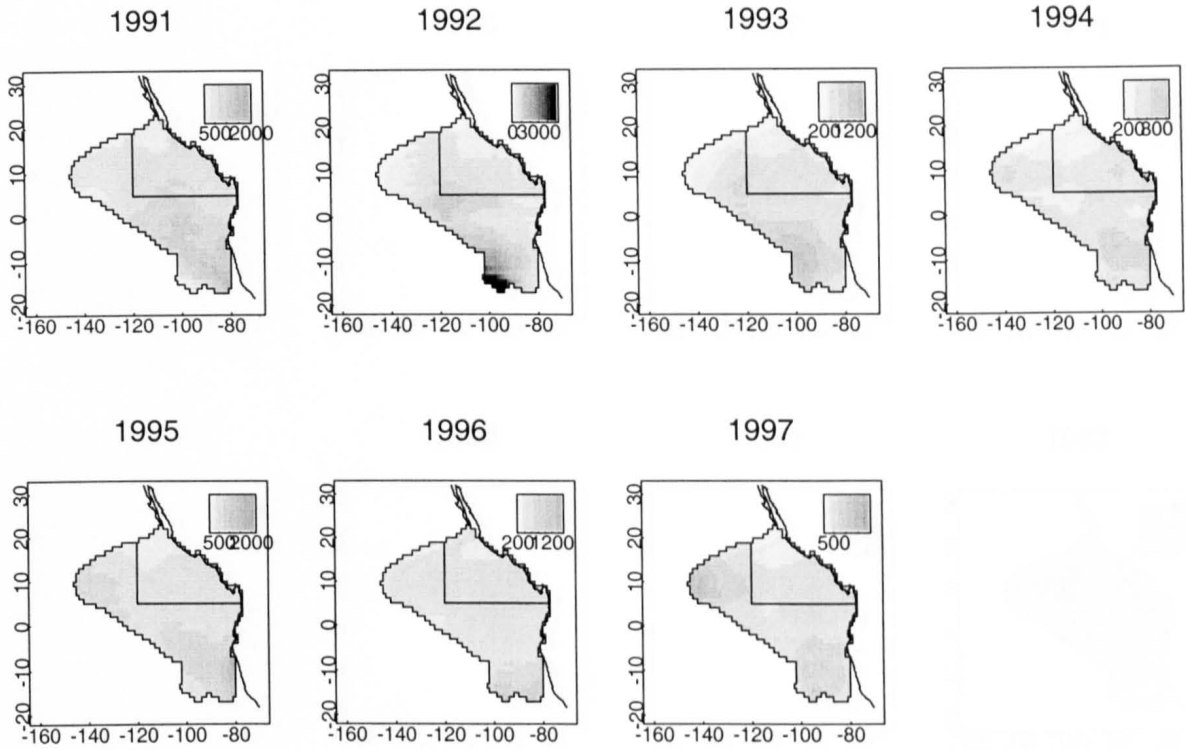


Figure 2.6: Predicted mean school size surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.[continued]

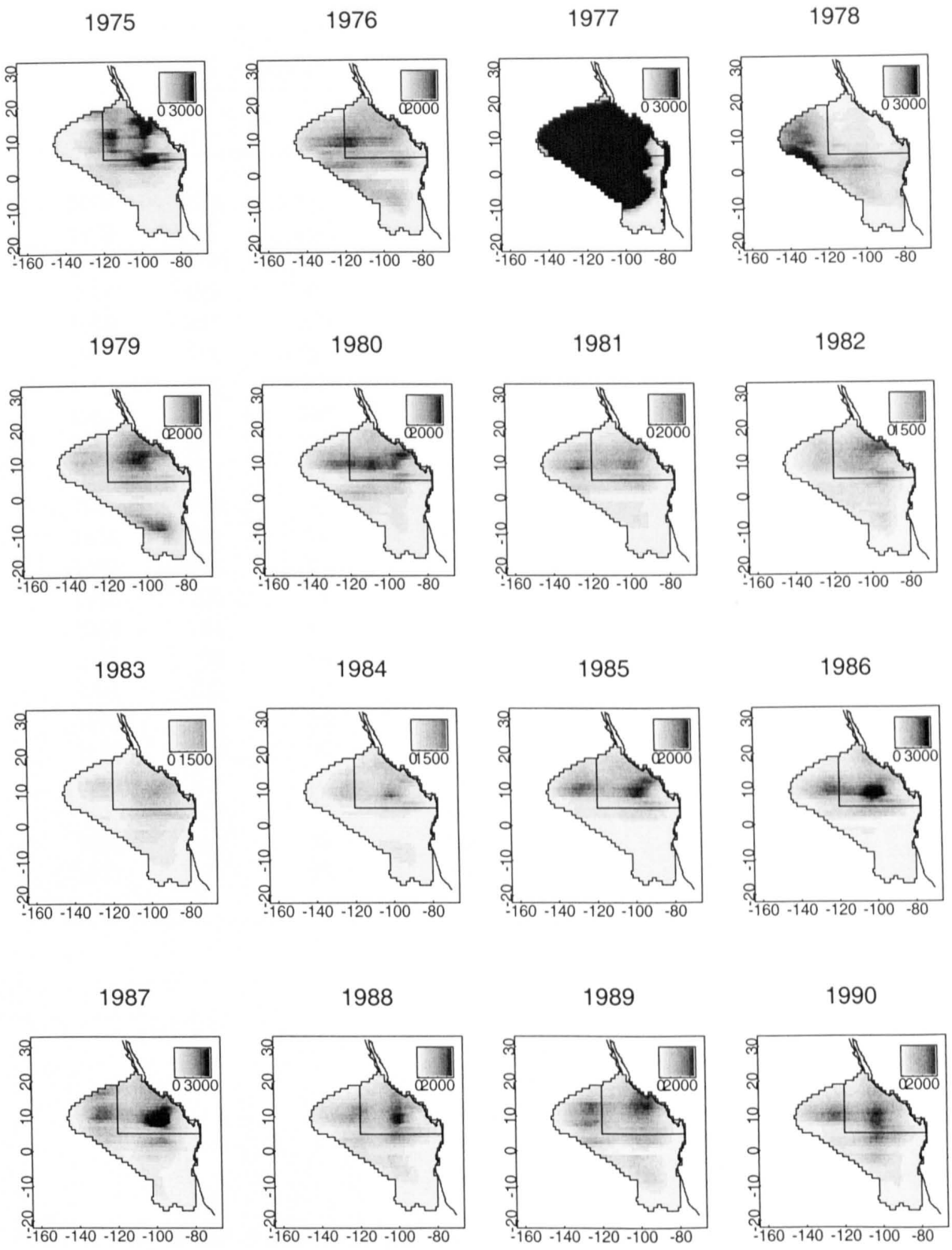


Figure 2.7: Predicted density surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.

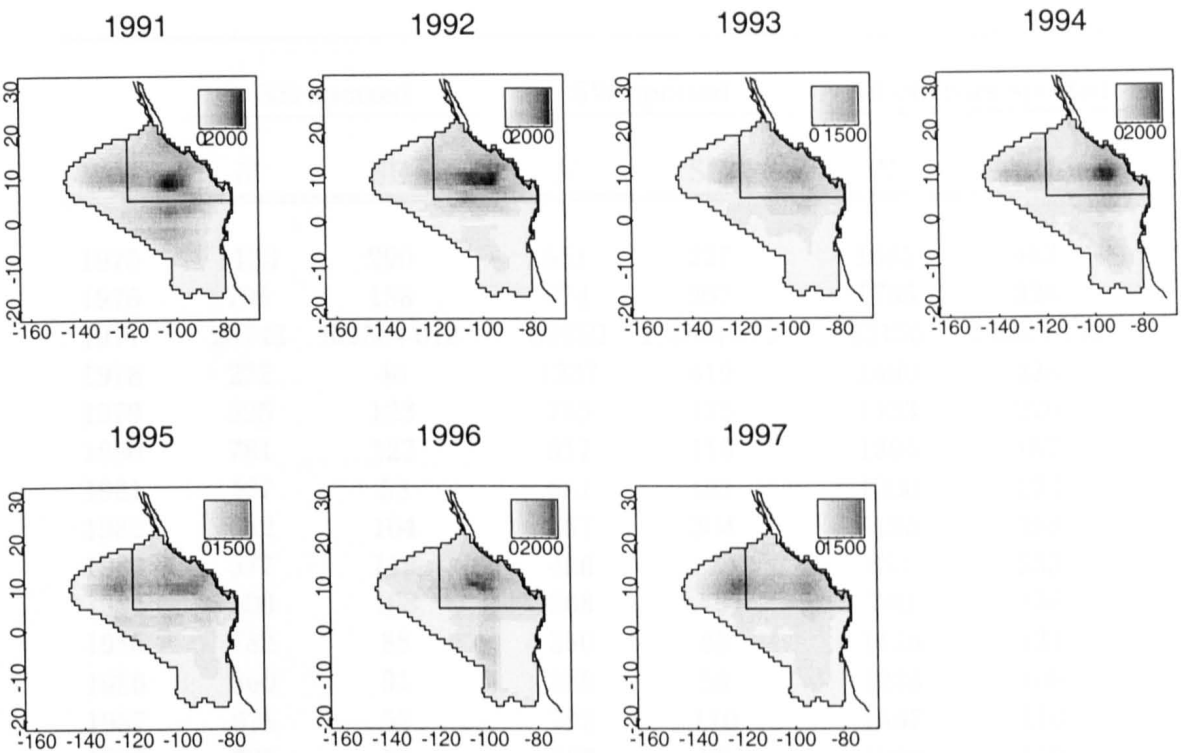
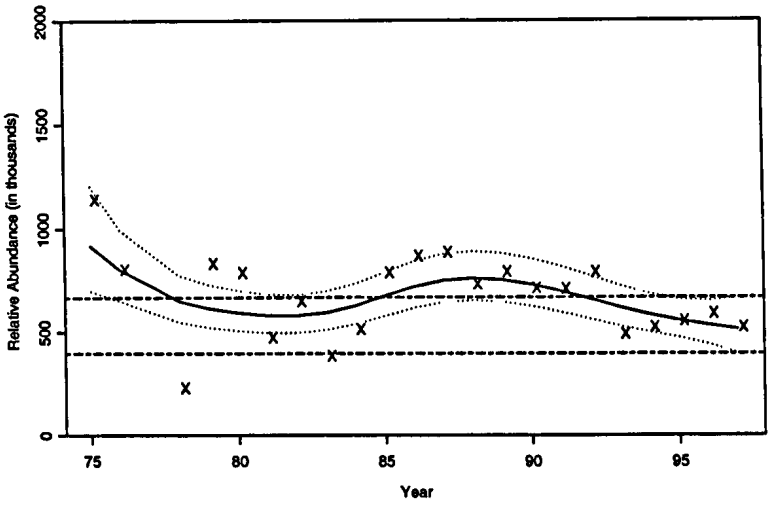


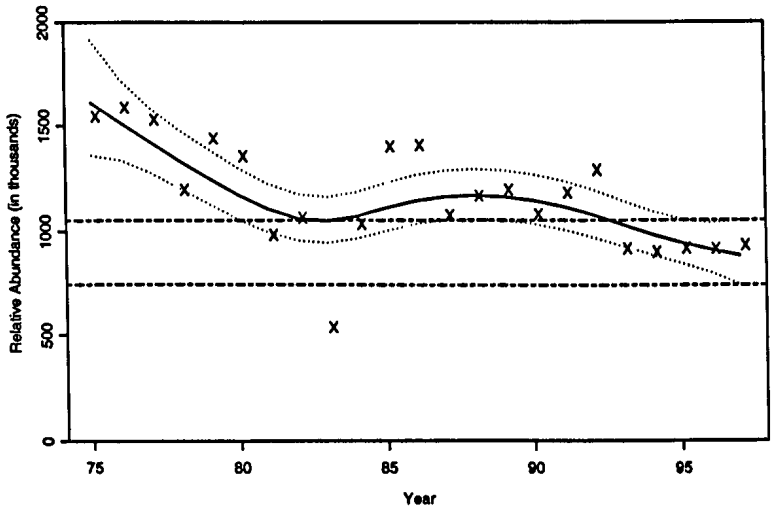
Figure 2.7: Predicted density surfaces for offshore spotted dolphins, averaged across months within each year, for the years 1975–1997.[continued]

	NE spotted		SW spotted		All offshore spotted	
Year	$\hat{N}$	SE	$\hat{N}$	SE	$\hat{N}$	SE
1975	1133	290	551	227	1685	483
1976	796	188	974	207	1785	328
1977	25743	5.95e+018	58250	2.47e+019	85155	2.85e+019
1978	222	40	1237	410	1460	428
1979	825	123	755	175	1433	226
1980	781	123	617	115	1394	187
1981	467	58	681	151	1000	124
1982	642	104	667	304	1195	298
1983	377	119	406	128	831	233
1984	506	108	268	73	740	136
1985	782	88	390	68	1128	134
1986	860	91	389	58	1235	109
1987	878	52	793	110	1667	110
1988	725	70	567	82	1283	112
1989	784	68	751	89	1505	113
1990	706	49	629	71	1320	93
1991	702	49	725	85	1337	99
1992	784	56	348	41	1100	71
1993	483	30	299	30	775	48
1994	517	28	476	41	985	49
1995	547	31	499	54	1066	58
1996	582	39	346	66	932	92
1997	517	25	466	45	973	51

Table 2.4: Estimates of relative abundance and its standard error for northeastern offshore spotted dolphins, southern-western offshore spotted dolphins, and all stocks of offshore spotted dolphins combined, for the years 1975 – 1997.

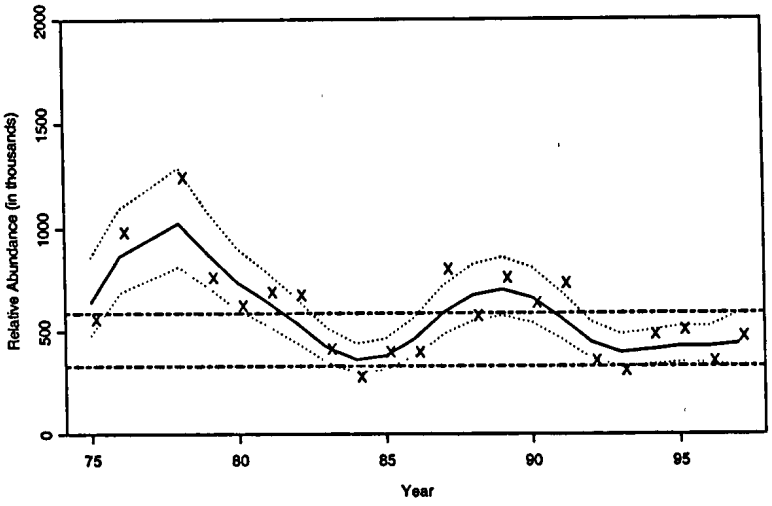


(a) Spatio-temporal modelling

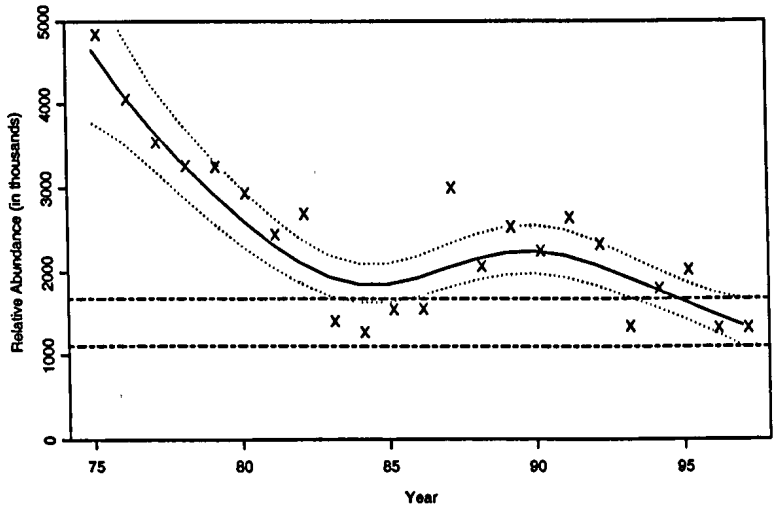


(b) Post-stratification method

Figure 2.8: Trends in relative abundance for the northeastern stock of offshore spotted dolphins for the years 1975–1997, obtained using (a) the spatio-temporal modelling approach and (b) using the post-stratification method. Dashed lines indicate 85% confidence intervals.



(a) Spatio-temporal modelling



(b) Post-stratification method

Figure 2.9: Trends in relative abundance for the southern-western stock of offshore spotted dolphins for the years 1975–1997, obtained using (a) the spatio-temporal modelling approach and (b) using the post-stratification method. Dashed lines indicate 85% confidence intervals.

2.6 Discussion

The spatio-temporal models presented in this Chapter provide a means of obtaining estimates of density and abundance from line transect data obtained in the absence of a proper survey design. However, although the models do not require large sample sizes, they are constrained by the need for good spatial and temporal coverage of the survey region. The spatio-temporal models fitted to offshore spotted dolphin sightings data appeared to capture well the observed distribution of encounter rate and mean cluster size within the northeastern offshore spotted dolphin stock boundary. In most years there is good spatial and temporal coverage both within and immediately outside the stock area, and hence the models do not extrapolate beyond the range of the data. In contrast, search effort is generally sparse within the boundaries of the southern-western stock of offshore spotted dolphins, and some of the predicted values near the outer range of the survey region may not be reliable. Structural zeros were useful in preventing the models from predicting unrealistically high values near the outer stock boundary. However, simulations are required to determine exactly how the models behave under those circumstances.

Provided good spatial and temporal coverage of the survey region, an advantage of the spatio-temporal modelling approach, in comparison with some design-based methods, is that it allows a more detailed examination of the predicted distribution of the population of interest. This in turn may help our understanding of ecological and behavioural characteristics of that population. In addition, it provides a means of assessing some of the assumptions we make about the population. Take the northeastern stock of offshore spotted dolphins as an example. Estimates of relative abundance for that stock are based on a set of fixed stock boundaries, determined based on morphological, biological and distributional characteristics (Dizon *et al.* 1994). In practice it is likely that such boundaries vary annually, or even monthly, according to changes in environmental conditions and prey availability. Indeed our models show variability in the predicted density surface for that stock between years, which suggests that some amount of movement across the stock boundaries might be taking place. If that is the case, then much of the annual fluctuations in dolphin relative abundance will be a reflection of such movement, rather than of an underlying trend in the population.

One of the motivations for the development of spatio-temporal models for line transect data was to try to improve the accuracy and precision of estimates of ETP dolphin relative abundance based on tuna vessel observer data. For the stocks of offshore spotted dolphin,

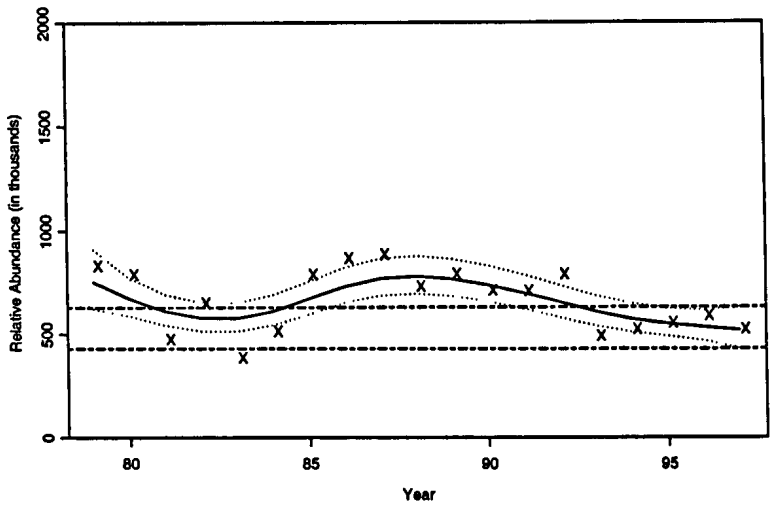
the CVs for the estimates based on the spatio-temporal models were generally smaller than those obtained following the procedures of Buckland and Anganuzzi (1988*b*) and Anganuzzi and Buckland (1989), with the exception of estimates for the 1970s, which were greater than those obtained using the post-stratification procedure. In either case, nevertheless, the differences were small.

However, point estimates of offshore spotted dolphin relative abundance obtained from the spatio-temporal models were consistently lower than those obtained using the traditional approach. This is likely a result of the different way in which the methods generate a predicted surface over areas where there are no observations. Search effort near the southern boundary of the survey region is generally very limited, especially in the early years. Under the traditional approach, the smoothing procedure will generate values for that region which are somewhat similar to those from the nearest area where there is survey effort. But the nearest area where there is effort is usually located further north, where dolphin densities are greater. In contrast, the spatio-temporal models are being influenced by the structural zeros, so that the further south you go, the lower are the encounter rates and mean cluster size; hence the lower density values. If this assumption is correct, and the available data appears to support it (e.g. Dizon *et al.* 1992, 1994), then the predicted surfaces from the spatio-temporal models are likely more realistic than those obtained based on the traditional approach.

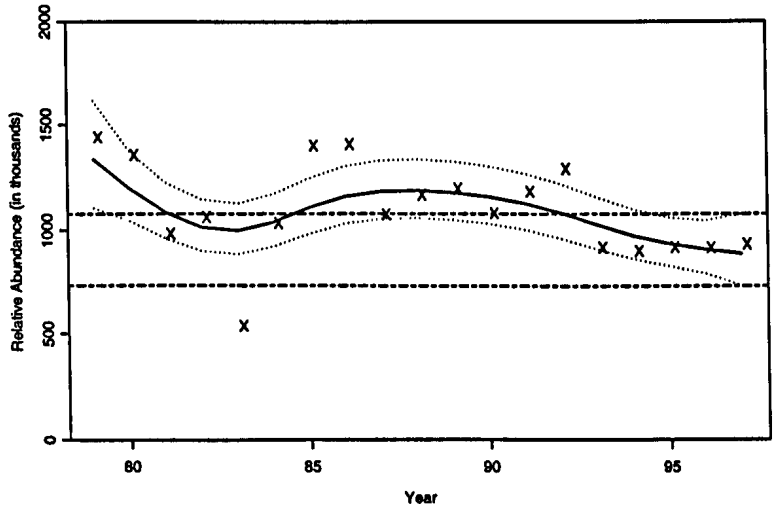
Nonetheless, trends in relative abundance for the northeastern stock of offshore spotted dolphin were comparable between the two methods, with the spatial-temporal modelling procedure yielding a slightly smoother trend. Both methods suggest that the population has been relatively stable since the early 1980s. Although a decline is apparent during the 1990s, it is not significant at the 5% level. Given the observed decline around the strong El Niño year of 1983, it is possible that the same pattern is being repeated in the late 1990s as a result of the strong 1997–1998 El Niño event. The spatio-temporal models also gave smoother trends than the post-stratification method for the southern-western stock of offshore spotted dolphins (note that the scale in the y-axis of the plots in Figure 2.9 differ; if we were to use the same scale in both plots, we would almost obtain flat trend estimates for the spatio-temporal models). Both methods suggest a decline in relative abundance since the 1970s, with an apparently cyclic variation which coincides with the occurrence of strong El Niño events. For the late 1990s, nonetheless, the spatio-temporal models suggest a much more stable trend than the post-stratification method. According to results from the spatio-temporal models, there appears to be no significant trends in southern-western offshore spotted dolphin relative abundance, except for an

apparent decline between 1978 and 1984. However, a word of caution is required for the interpretation of trends in relative abundance for the 1970s. Even for the northeastern stock of offshore spotted dolphins, which is located within an area with good coverage, some of the estimates of relative abundance for the early years are very large. While this may lead to the interpretation that there was a massive decline in abundance during those years, data quality was generally poor (*cf.* Lennert-Cody *et al.* In prep) and effort was limited, and hence it is likely that such high estimates are a reflection of the latter. This possibility is even more compelling for the southern-western stock of offshore spotted dolphins: while the post-stratification procedure suggests a decline from about 5 million animals in 1975 to less than 2 million in 1984, something which may not even be possible to occur from a biological point of view, the spatio-temporal models show that population almost doubling its size and then returning to approximately the same level within a six year period (1975–1980). We do not claim that a decline has not occurred since the 1970s; however, given issues of data quality and coverage, we believe that the magnitude of such decline during the late 1970s cannot be quantified using the tuna vessel observer data. Figures 2.10 and 2.11 show trends in northeastern and southern-western offshore spotted dolphin relative abundance for 1979–1997. For the northeastern stock, the spatio-temporal models indicate a significant decline since the late 1980s. However, estimates for the early years from both the spatio-temporal models and the post-stratification method are not significantly different from the most recent estimates. Trends for the southern-western stock based on the spatio-temporal models are considerably smoother than those given by the post-stratification method, and do not indicate any significant differences in relative abundance for that stock over time.

A number of issues remain to be addressed. The first is the need for simulation studies to assess how the models behave when required to extrapolate over areas of the survey region where data are scarce. A well known feature of cubic smoothing splines is that they are linear beyond the range of the data (Hastie and Tibshirani 1990). This implies that linear relationships between the response and the models' covariates are applied beyond the range of the observed values, which may lead to unreliable estimates. In general extrapolation beyond the range of the data should be avoided. However, in the case of the tuna vessel observer data, search effort is often small or non-existent near the outer boundaries of some of the stocks for which estimates of relative abundance are required for management purposes, and extrapolation provides the only means of obtaining such estimates. The use of structural zeros allows the incorporation of prior knowledge about the distribution of dolphin populations within the modelling framework. If we can understand the models'

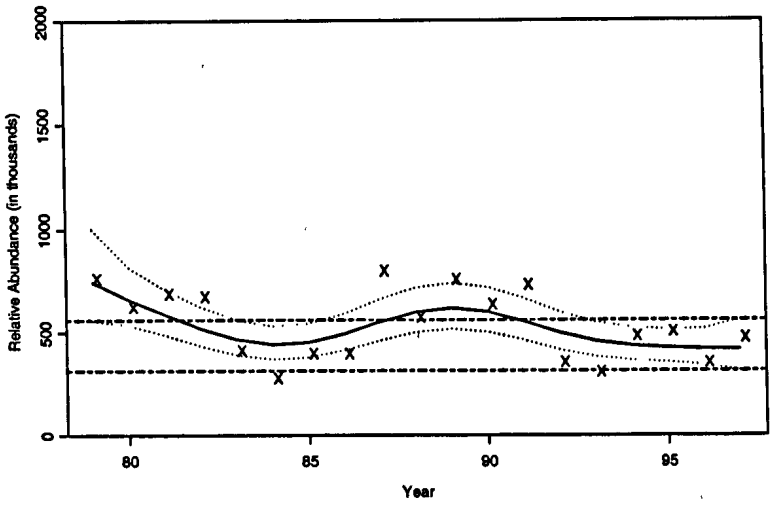


(a) Spatio-temporal modelling

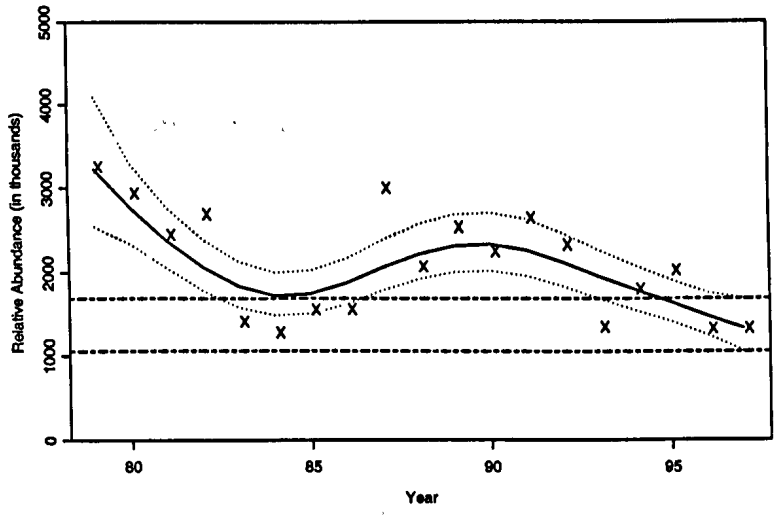


(b) Post-stratification method

Figure 2.10: Trends in relative abundance for the northeastern stock of offshore spotted dolphins for the years 1979–1997, obtained using (a) the spatio-temporal modelling approach and (b) using the post-stratification method. Dashed lines indicate 85% confidence intervals.



(a) Spatio-temporal modelling



(b) Post-stratification method

Figure 2.11: Trends in relative abundance for the southern-western stock of offshore spotted dolphins for the years 1979–1997, obtained using (a) the spatio-temporal modelling approach and (b) using the post-stratification method. Dashed lines indicate 85% confidence intervals.

behaviour when extrapolating over such areas, we can make more effective use of structural zeros to obtain reliable estimates for those dolphin stocks.

A second issue is the incorporation of the variability in the environmental covariates included in the models. Currently the grid of values of the oceanographic covariates is obtained based on a smoothing procedure applied to oceanographic observations from the Pacific Ocean (Fiedler 1992). These values are then effectively treated as constants in the models, in the same way that locational covariates (*i.e.* latitude and longitude) are used. In practice, however, the oceanographic covariates are random variables, each belonging to some distribution. One way of incorporating this variability in the oceanographic covariates is by using mixed models. In mixed models some of the covariates are treated as "known" (e.g. latitude and longitude) whereas the variance associated with the remaining covariates is directly incorporated into the estimation procedure. However, while the theoretical basis for fitting mixed models within a GLM framework is well developed (e.g. McCullagh and Nelder 1989), they are not yet available for GAMs.

2.7 Appendices

2.7.1 Derivation of cubic smoothing splines

The derivation below is taken primarily from Green and Silverman (1994).

2.7.1.1 Definitions and notation

Let $t_1, \ldots, t_n$ denote a set of real numbers in the interval $[a, b]$, where $a < t_1 < \ldots < t_n < b$. A function g is defined to be a cubic smoothing spline in the interval $[a, b]$ if:

1. g is a cubic polynomial on each interval $(a, t_1), (t_1, t_2), \ldots, (t_n, b)$; and
2. the piecewise polynomials join smoothly at the points (also referred to as knots) $t_1, \ldots, t_n$, so that g, g' and g'' are continuous at each t_i, and thus on all of $[a, b]$.

A cubic smoothing spline g in the interval $[a, b]$ is a natural (cubic) smoothing spline if its second and third derivatives are equal to zero at a and b, so that g is linear in the intervals $[a, t_1]$ and $[t_n, b]$. These are called the natural boundary conditions.

For the remainder of this derivation let $g_i = g(t_i)$, the value of g at the ith knot, and let $\gamma_i = g''(t_i)$, the second derivative of g at the ith knot.

2.7.1.2 Constructing a cubic smoothing spline

Although g can be intuitively represented as:

$$g(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3$$

for $t_i \leq t \leq t_{i+1}$, it turns out that it is easier to construct a cubic smoothing spline using the value of g and its second derivative within an interval. Thus given the interval $[t_L, t_R]$, let $g(t_L) = g_L$, $g(t_R) = g_R$, $g''(t_L^+) = \gamma_L$, and $g''(t_R^-) = \gamma_R$.

Since g is cubic, it follows that its second derivative is linear on $[t_L, t_R]$, so that, using standard linear algebra and defining $h = t_R - t_L$ we have:

$$g''(t) = \frac{(t - t_L)\gamma_R + (t_R - t)\gamma_L}{h} \tag{2.15}$$

and

$$g'''(t) = \frac{\gamma_R - \gamma_L}{h}. \tag{2.16}$$

Green and Silverman (1994) state that:

$$\begin{aligned} g(t) &= \frac{(t - t_L)g_R + (t_R - t)g_L}{h} \\ &\quad -\frac{1}{6}(t - t_L)(t_R - t)\left\{\left(1 + \frac{t - t_L}{h}\right)\gamma_R + \left(1 + \frac{t_R - t}{h}\right)\gamma_L\right\}. \end{aligned} \tag{2.17}$$

A proof of this statement can be obtained by showing that the four required conditions are met - namely that $g(t_R) = g_R$, $g(t_L) = g_L$, $g''(t_L^+) = \gamma_L$ and $g''(t_R^-) = \gamma_R$. The first two conditions can be easily shown to hold by substituting t for t_R and t_L in the previous equation. Thus for $t = t_R$:

$$g(t_R) = \frac{(t_R - t_L)\, g_R + (t_R - t_R)\, g_L}{h} - 0$$

$$= g_R$$

and, for $t - t_L$:

$$g(t_L) = \frac{(t_L - t_L)\, g_R + (t_R - t_L)\, g_L}{h} - 0$$

$$= g_L.$$

Now it remains to show that the second derivative of expression (2.17), evaluated at $t = t_R$ and $t = t_L$, corresponds to γ_R and γ_L, respectively. Hence, ignoring the first term in that equation (which is linear, and hence will vanish once it has been differentiated twice), we have:

$$g'(t) = -\frac{1}{6}\ (-2t + t_R + t_L)\left\{\left(1 + \frac{t - t_L}{h}\right)\gamma_R + \left(1 + \frac{t_R - t}{h}\right)\gamma_L\right\}$$

$$-\frac{1}{6}\left(-t^2 + tt_R + tt_L + t_R t_L\right)\left\{\frac{\gamma_R - \gamma_L}{h}\right\}$$

and

$$g''(t) = \frac{1}{3}\ \left\{\left(1 + \frac{t - t_L}{h}\right)\gamma_R + \left(1 + \frac{t_R - t}{h}\right)\gamma_L\right\}$$

$$-\frac{1}{6}(-2t + t_R + t_L)\left\{\frac{\gamma_R - \gamma_L}{h}\right\}$$

$$-\frac{1}{6}(-2t + t_R + t_L)\left\{\frac{\gamma_R - \gamma_L}{h}\right\}$$

$$-\frac{1}{6}\left(-t^2 + tt_R + tt_L - t_R t_L\right) \cdot 0$$

$$= \frac{1}{3} \left\{ \left(1 + \frac{t - t_L}{h}\right) \gamma_R + \left(1 + \frac{t_R - t}{h}\right) \gamma_L \right\}$$

$$-\frac{1}{3}(-2t + t_R + t_L) \left\{ \frac{\gamma_R - \gamma_L}{h} \right\}.$$

Thus, for $t = t_R$ we have:

$$g''(t_R) = \frac{1}{3} \left\{ \left(1 + \frac{t_R - t_L}{h}\right) \gamma_R + \left(1 + \frac{t_R - t_R}{h}\right) \gamma_L \right\}$$

$$+\frac{1}{3} \left\{ \frac{t_R - t_L}{h} \right\} \{\gamma_R - \gamma_L\}$$

$$= \frac{1}{3} \{2\gamma_R + \gamma_L + \gamma_R - \gamma_L\}$$

$$= \gamma_R$$

and, for $t = t_L$:

$$g''(t_l) = \frac{1}{3} \left\{ \left(1 + \frac{t_L - t_L}{h}\right) \gamma_R + \left(1 + \frac{t_R - t_L}{h}\right) \gamma_L \right\}$$

$$-\frac{1}{3} \left\{ \frac{t_R - t_L}{h} \right\} \{\gamma_R - \gamma_L\}$$

$$= \frac{1}{3} \{\gamma_R + 2\gamma_L - \gamma_R + \gamma_L\}$$

$$= \gamma_L,$$

completing the proof.//

Expression (2.17) can be used to obtain values of g within each interval $[t_i, t_{i+1}]$ on $[t_1, t_n]$. This can be shown by defining $h_i = t_{i+1} - t_i$, for $i = 1, \ldots, n-1$. Then:

$$g(t) = \frac{(t-t_i)g_{i+1} + (t_{i+1}-t)g_i}{h_i} - \frac{1}{6}(t-t_i)(t_{i+1}-t)\left\{\left(1+\frac{t-t_i}{h_i}\right)\gamma_{i+1} + \left(1+\frac{t_{i+1}-t}{h_i}\right)\gamma_i\right\} \tag{2.18}$$

for $t_i \leq t \leq t_{i+1}$. To obtain values for g when $t \leq t_1$ or $t \geq t_n$ (i.e. outside the range of the knots), recall that from the boundary conditions we know that g must be linear on the extreme intervals $[a, t_1]$ and $[t_n, b]$. From expression (2.17) we have:

$$g'(t_L^+) = \frac{g_R - g_L}{h} - \frac{1}{6}h(2\gamma_L + \gamma_R) \tag{2.19}$$

and

$$g'(t_R^-) = \frac{g_R - g_L}{h} + \frac{1}{6}h(\gamma_L + 2\gamma_R). \tag{2.20}$$

Hence the derivatives of g at t_1 and t_n are given by:

$$g'(t_1) = \frac{g_2 - g_1}{t_2 - t_1} - \frac{1}{6}(t_2 - t_1)\gamma_R \tag{2.21}$$

and

$$g'(t_n) = \frac{g_n - g_{n-1}}{t_n - t_{n-1}} + \frac{1}{6}(t_n - t_{n-1})\gamma_{n-1}. \tag{2.22}$$

A value for g' at $t = t_0$, where $t_0 < t_1$, is then obtained by substituting t_0 for t_1 in expression (2.21) and solving for g_0:

$$g'(t_0) = \frac{g_1 - g_0}{t_1 - t_0} - \frac{1}{6}(t_1 - t_0)\gamma_1.$$

Noting that from the boundary conditions we have that $\gamma_1 = 0$ and $g'(t_0) = g'(t_1)$, then:

$$g'(t_1)(t_1 - t_0) = g_1 - g_0$$

so that:

$$g_0 = g_1 - (t_1 - t_0)g'(t_1).$$

For $t = t_{n+1}$ we have:

$$g'(t_{n+1}) = \frac{g_{n+1} - g_n}{t_{n+1} - t_n} - \frac{1}{6}(t_{n+1} - t_n)\gamma_n.$$

Noting that $\gamma_n = 0$ and $g'(t_{n+1}) = g'(t_n)$, then:

$$g'(t_{n+1})(t_{n+1} - t_n) = g_{n+1} - g_n$$

and so:

$$g_{n+1} = g_n + (t_{n+1} - t_n)g'(t_n).$$

We can now evaluate the cubic smoothing spline g at each interval $(t_0, t_1), \ldots, (t_n, t_{n+1})$. It remains for us to ensure that g is continuous at all knots; that is, $g'(t_i^+)$ must equal $g'(t_i^-)$. Substituting t_i^+ and t_i^- in expressions (2.19) and (2.20), and equating the two, we have:

$$\frac{g_i - g_{i-1}}{h_{i-1}} + \frac{1}{6}h_{i-1}(\gamma_{i-1} + 2\gamma_i) = \frac{g_{i+1} - g_i}{h_i} - \frac{1}{6}h_i(2\gamma_i + \gamma_{i+1})$$

and so:

$$\begin{aligned} g_{i+1}h_i^{-1} + g_i(-h_i^{-1}) + g_i(-h_{i-1}^{-1}) + g_{i-1}h_{i-1}^{-1} &= \frac{1}{6}\gamma_{i+1}h_i + \frac{1}{3}\gamma_i h_i + \frac{1}{6}\gamma_{i-1}h_{i-1} + \frac{1}{3}\gamma_i h_{i-1} \\ g_{i+1}h_i^{-1} + g_i((-h_i^{-1})(-h_{i-1}^{-1})) + g_{i-1}h_{i-1}^{-1} &= \frac{1}{6}\gamma_{i-1}h_{i-1} + \frac{1}{3}\gamma_i(h_i h_{i-1}) + \frac{1}{6}\gamma_{i+1}h_i. \end{aligned} \tag{2.23}$$

Using the above equation we can then construct the matrices required to solve for γ (see p.12 in Green and Silverman (1994) for details).

2.7.2 Cubic smoothing splines as the solution to the PRSS

We use the definitions and notations from Appendix 2.7.1. In addition, let $y_1, \ldots, y_n$ denote observations, and let $S(g)$ be the PRSS, given by:

$$S(g) = \sum_{i=1}^{n} \{y_i - g(t_i)\}^2 + \lambda \int_a^b \left\{g''(x)\right\}^2 dx \tag{2.24}$$

and described in section 2.3. Denote by $\hat{g}$ the estimated curve which minimises $S(g)$. Here we present results given by Green and Silverman (1994), which show that $\hat{g}$ is a natural cubic spline with knots at the points t_i.

Given a function $\tilde{g}(t_i) = z_i$, define a function $h = \tilde{g} - g$. Since both $\tilde{g}$ and g interpolate the values z_i, then $h = 0$ at all points t_i. We need now to solve the integral $\int_a^b g''(t)\, h''(t)\, dt$, a result which will be needed later. Integrating by parts gives:

$$\begin{aligned} \int_a^b g''(t)\, h''(t)\, dt &= g''(t)\, h'(t)\Big|_a^b - \int_a^b h'(t)\, g'''(t)\, dt \\ &= g''(b)\, h'(b) - g''(a)\, h'(a) - \sum_{j=1}^{n-1} g'''(t_j^+) \int_{t_j}^{t_{j+1}} h'(t)\, dt. \end{aligned}$$

The first two terms will vanish as, according to the definition of a natural cubic spline, $g''(a) = g''(b) = 0$. Further, since $g'''(t)$ is zero on the intervals (a, t_1) and (t_n, b), and is constant on (t_j, t_{j+1}) with value $g'''(t_j^+)$, we obtain the last term on the right hand side of the previous equation. We then have:

$$\begin{aligned} \int_a^b g''(t)\, h''(t)\, dt &= -\sum_{j=1}^{n-1} g'''(t_j^+) \int_{t_j}^{t_{j+1}} h'(t)\, dt \\ &= -\sum_{j=1}^{n-1} g'''(t_j^+) \{h(t_{j+1}) - h(t_j)\} \qquad (2.25) \\ &= 0 \end{aligned}$$

since $h(\cdot) = 0$. The above result can then be applied in the derivation below:

$$\begin{aligned}\int_a^b \left\{\tilde{g}''(t)\right\}^2 dt &= \int_a^b \left(g''(t) + h''(t)\right)^2 dt \\ &= \int_a^b \left\{g''(t)\right\}^2 dt + 2\int_a^b g''(t)\,h''(t)\,dt + \int_a^b \left\{h''(t)\right\}^2 dt \qquad (2.26) \\ &= \int_a^b \left\{g''(t)\right\}^2 dt + \int_a^b \left\{h''(t)\right\}^2 dt \geq \int_a^b \left\{g''(t)\right\}^2 dt.\end{aligned}$$

which shows that $\int_a^b \{h''(t)\}^2\,dt$ must be equal to zero for the equality to hold, so that h is linear on $[a,b]$. However, as by definition $h = 0$ at all t_i, then g and $\tilde{g}$ are the same function - that is, a natural cubic spline.

Having shown that the function that minimises the penalty term is a natural cubic spline, it remains for us to show that the minimiser of $S(g)$ is also a natural cubic spline. To this end, let g be any curve which is not a natural cubic spline, and define $\overline{g}$ to be the natural cubic spline which interpolates the values of $g(t_i)$. Thus, by definition, $\overline{g}(t_i) = g(t_i)$ for all i, and it follows that $\sum_{i=1}^n \{y_i - \overline{g}(t_i)\}^2 = \sum_{i=1}^n \{y_i - g(t_i)\}^2$. Given the result from expression (2.26), then $\int\{\overline{g}''(t)\}^2\,dt < \int\{g''(t)\}^2\,dt$. As $\lambda > 0$, then it follows that $S(\overline{g}) < S(g)$; that is, the minimiser of $S(g)$ is a natural cubic spline.

2.7.3 Validating the use of AIC in the context of GAMs

AIC is a commonly used criterion to select amongst competing models, and is defined as:

$$\begin{aligned}\text{AIC} &= D + 2p \\ &= 2\sum_{i=1}^n \mathbb{E}\left[\log \pi(y_i;\theta) - \log \pi(y_i;\hat{\theta})\right] + 2p,\end{aligned} \qquad (2.27)$$

The first term in the expression above is the deviance, given by twice the log-likelihood of the true model (or a saturated model) minus the log-likelihood obtained from some contending model with estimated parameters $\hat{\theta}$. Here $y_i (i = 1,\ldots,n)$ denote iid observations from some distribution, and $\pi(y_i;\theta)$ specifies a model. The second term penalises for the number of parameters p in the contending model. Akaike (1973) shows that minimising

the deviance is equivalent to maximising the negative log-likelihood of the contending model; that is, $D = -2\sum_{i=1}^{n} \log \pi(y_i; \hat{\theta})$.

Two underlying assumptions implicit in the use of AIC are that:

1. parameter estimates are obtained via maximum likelihood, as AIC is based on the Kullback-Leibler (Kullback and Leibler 1951) measure of distance; and
2. the true model (or family of models) is amongst the contending models.

Assumption (1) is violated in the context of GAMs, where the smooth functions are obtained via an iterative procedure (the 'backfitting algorithm'; Hastie and Tibshirani 1990) based on the minimisation of the penalised residual sum of squares. Assumption (2) is usually unrealistic, as we have no means of asserting whether the true model is one of the contending models.

Moody (1991, 1992) and Murata *et al.* (1994) derived generalisations of AIC in the context of neural networks which allow for any function to be used in place of the likelihood, and where the true model is not necessarily included amongst the contending models. Here we show how their generalisation easily applies to GAMs, and hence justifies the use of AIC in that context. The derivation given below is taken primarily from Ripley (1996).

Let $\mathbf{y} = y_1, \ldots, y_n$ denote iid observations with pdf $\pi(\mathbf{y})$, and let $\pi(\mathbf{y}; \theta)$ be a parametric model fitted to the data and where θ is a q-dimensional parameter vector. Denote by $\hat{\theta}$ the estimate of θ which maximises the log-likelihood $\ell(\theta; \mathbf{y}) = \sum_{i=1}^{n} \log \pi(y_i; \theta)$ with respect to θ. Also, suppose that this function has a unique maximum at $\theta = \theta_0$, where θ_0 is not necessarily assumed to correspond to the true parameter value.

Ripley (1996) shows that $n \times D$, the latter being the expected deviance from expression (2.27), corresponds to:

$$n \times D = 2 \sum_{i=1}^{n} \mathbb{E}\left[\log \pi(y_i; \theta_0) - \log \pi(y_i; \hat{\theta})\right] + 2p^* + O(1/\sqrt{n}), \tag{2.28}$$

where $p^* = \text{trace}[KJ^{-1}]$, with

$$J = -\mathbb{E}\left[\frac{\partial^2 \log \pi(\mathbf{y}; \theta_0)}{\partial \theta \partial \theta^T}\right] \tag{2.29}$$

and

$$K = \left[\text{var}\left(\frac{\partial \log \pi(\mathbf{y};\theta_0)}{\partial \theta}\right)\right], \tag{2.30}$$

where J is the negative expectation of the Hessian matrix and K corresponds to Fisher's information matrix. The error in expression (2.28) is of the order $1/\sqrt{n}$. If the parametric model fitted to the data contains the true model, then $p^* = p$, the number of parameters in the model.

It can be shown (Ripley 1996) that $\hat{\theta}$ converges to θ_0 without the need to assume that the parametric model $\pi(\mathbf{y};\hat{\theta})$ belongs to the family which contains the true model. Also, by changing the definition of J and K, the above results hold for any function $\psi(\cdot)$ other than the likelihood, and which gives parameter estimates $\hat{\theta}$ that minimise $\sum_{i=1}^{n} \psi(\theta; y_i)$. Thus, in the context of GAMs, replace the log-likelihood by the PRSS from expression (2.2). For notational simplicity we rewrite expression (2.2) as $\psi(\theta; y_i) = R(y_i;\theta) + \lambda C(\theta)$. Substituting $\ell(\cdot)$ for $\psi(\cdot)$ gives the Network Information Criterion (NIC) of Murata *et al.* (1994):

$$\begin{aligned} \text{NIC} &= 2\sum_{i=1}^{n}\left\{(R(y_i;\theta) + \lambda C(\theta)) - (R(y_i;\hat{\theta}) + \lambda C(\hat{\theta}))\right\} + 2p^* \\ &= D + 2p^* \end{aligned} \tag{2.31}$$

where the first term gives a measure of the model fit (deviance), $p^* = \text{trace}\,[KJ^{-1}]$, and

$$J = \mathbb{E}\left[\frac{\partial^2 \psi(\theta_0;\mathbf{y})}{\partial \theta \partial \theta^T}\right] \tag{2.32}$$

and

$$K = \left[\text{var}\left(\frac{\partial \psi(\theta_0;\mathbf{y})}{\partial \theta}\right)\right]. \tag{2.33}$$

As the integral from expression (2.2) does not depend on the data, it does not contribute to the variance and hence can be omitted from expression (2.33). Thus we have:

$$K = \left[\text{var}\left(\frac{\partial R(\mathbf{y};\theta_0)}{\partial\theta}\right)\right]. \tag{2.34}$$

Thus NIC provides a generalisation of AIC which can be applied to any model fitted to data based on any objective function other than the likelihood, and which does not rely on the assumption that the contending model belongs to the family of models which includes the true parameter value θ_0.

The question then is how to estimate J and K, and hence p^*. To obtain an estimate of J we can simply replace θ_0 by their estimates $\hat{\theta}$; this is the approach of Moody (1991, 1992). Rewriting expression (2.2) in matrix notation (*cf.* Hastie and Tibshirani 1990, p. 29) gives:

$$\begin{aligned}\psi(\theta;\mathbf{y}) &= (\mathbf{y}-\mathbf{f})^T(\mathbf{y}-\mathbf{f}) + \lambda\mathbf{f}^T\Omega\mathbf{f} \\ &= \mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{f} + \mathbf{f}^T\mathbf{f} + \lambda\mathbf{f}^T\Omega\mathbf{f}.\end{aligned} \tag{2.35}$$

Then substituting the above expression into expression (2.32) yields:

$$\begin{aligned}\hat{J} &= \frac{\partial^2\mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{f} + \mathbf{f}^T\mathbf{f} + \lambda\mathbf{f}^T\Omega\mathbf{f}}{\partial\theta\partial\theta^T} \\ &= \frac{\partial 2(\mathbf{f}-\mathbf{y}+\lambda\Omega\mathbf{f})}{\partial\theta} \\ &= 2(I+\lambda\Omega).\end{aligned} \tag{2.36}$$

To estimate K, define $\delta(\mathbf{y}) = \partial R(\mathbf{y};\hat{\theta})/\partial\theta$. Then a Taylor's expansion of $\delta(\mathbf{y})$ about $\mathbf{y}$ gives:

$$\delta(\mathbf{y}) \approx \delta(\mathbf{y}^*) + \frac{\partial\delta(\mathbf{y})}{\partial\mathbf{y}}(\mathbf{y}-\mathbf{y}^*). \tag{2.37}$$

The first term in the above expression is a constant. Now, recall that if a vector of observations $\mathbf{y}$ has variance-covariance matrix $[\text{var}(\mathbf{y})] = I\sigma^2$, then from standard results in linear regression (*cf.* Seber 1977, p. 11) a function of $\mathbf{y}$, say $\mathbf{U} = \mathbf{A}\mathbf{y}$, will have variance-covariance matrix $[\text{var}(\mathbf{U})] = \mathbf{A}\,[\text{var}(\mathbf{y})]\,\mathbf{A}^T$. Hence:

$$[\widehat{\text{var}}\{\delta(\mathbf{y})\}] \approx \left[\frac{\partial\delta(\mathbf{y})}{\partial\mathbf{y}}\right] I\sigma^2 \left[\frac{\partial\delta(\mathbf{y})}{\partial\mathbf{y}}\right]^T. \tag{2.38}$$

Substituting $\partial R(\mathbf{y};\hat{\theta})/\partial\theta$ for $\delta(\mathbf{y})$ gives:

$$\begin{aligned} \hat{K} &= \left[\widehat{\text{var}}\left\{\frac{\partial R(\mathbf{y};\hat{\theta})}{\partial\theta}\right\}\right] \\ &\approx \left[\frac{\partial^2 R(\mathbf{y};\hat{\theta})}{\partial\mathbf{y}\partial\theta}\right] I\sigma^2 \left[\frac{\partial^2 R(\mathbf{y};\hat{\theta})}{\partial\mathbf{y}\partial\theta}\right]^T, \end{aligned} \tag{2.39}$$

where $I\sigma^2$ is the variance of the $\mathbf{y}$ and

$$\begin{aligned} \left[\frac{\partial^2 R(\mathbf{y};\hat{\theta})}{\partial\mathbf{y}\partial\theta}\right] &= \frac{\partial^2 \mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{f} + \mathbf{f}^T\mathbf{f}}{\partial\mathbf{y}\partial\theta} \\ &= \frac{\partial 2(\mathbf{f}-\mathbf{y})}{\partial\mathbf{y}} \\ &= -2I. \end{aligned} \tag{2.40}$$

The trace of $\left[\hat{K}\hat{J}^{-1}\right]$ from expressions (2.36) and (2.39) is then given by:

$$\begin{aligned} \text{trace}\left[\hat{K}\hat{J}^{-1}\right] &= \text{trace}\left[2I\left(2\left(I+\lambda\Omega\right)\right)^{-1} I\sigma^2 2I\right] \\ &= \text{trace}\left[\left(I+\lambda\Omega\right)^{-1} 2\sigma^2 I\right] \\ &= 2\sigma^2 p'. \end{aligned} \tag{2.41}$$

The expression above corresponds to Moody's (1991, 1992) expression for computing the number of parameters in the model (*cf.* expression (14) in Moody 1992). By analogy with linear regression, it can then be shown that the estimate of p^* above should be replaced by $p^*/2\sigma^2$ (Simon Wood, University of St Andrews, Mathematical Institute, North Haugh, St Andrews, Scotland; personal communication). To see this, define $\tau(\beta;\mathbf{y})$ as the likelihood function to be maximised; that is:

$$\begin{aligned} \tau(\beta;\mathbf{y}) &= \frac{1}{2\sigma^2} \parallel \mathbf{y} - \mathbf{x}\beta \parallel^2 \\ &= \frac{1}{2\sigma^2} \left(\mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{x}^T\mathbf{y} + \beta^T\mathbf{x}^T\mathbf{x}\beta\right). \end{aligned} \tag{2.42}$$

Then

$$\begin{aligned} \hat{J} &= \frac{\partial^2(1/2\sigma^2)\left(\mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{x}^T\mathbf{y} + \beta^T\mathbf{x}^T\mathbf{x}\beta\right)}{\partial\beta\partial\beta^T} \\ &= \frac{\mathbf{x}^T\mathbf{x}}{\sigma^2} \end{aligned} \tag{2.43}$$

and

$$\begin{aligned} \left[\frac{\partial^2\tau(\beta;\mathbf{y})}{\partial\mathbf{y}\partial\beta}\right] &= \frac{\partial^2(1/2\sigma^2)\left(\mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{x}^T\mathbf{y} + \beta^T\mathbf{x}^T\mathbf{x}\beta\right)}{\partial\mathbf{y}\partial\beta^T} \\ &= -\frac{\mathbf{x}}{\sigma^2} \end{aligned} \tag{2.44}$$

so that

$$\hat{K} = \frac{\mathbf{x}^T\mathbf{x}}{\sigma^2}. \tag{2.45}$$

Then

$$\begin{aligned} \text{trace}\left[KJ^{-1}\right] &= \text{trace}\left[\frac{\mathbf{x}^T\mathbf{x}}{\sigma^2}\left(\frac{\mathbf{x}^T\mathbf{x}}{\sigma^2}\right)^{-1}\right] \\ &= p \end{aligned} \tag{2.46}$$

where p corresponds to the number of parameters in the model. If an objective function other than the likelihood is used, say the RSS, then following the derivation above we have trace $[KJ^{-1}] = p/2\sigma^2$. Hence, in the context of GAMs, where the objective function is the PRSS, we replace p^* by $p^*/2\sigma^2$. This corresponds to trace $\left[(I + \lambda\Omega)^{-1}\right]$, which is exactly how the *df* of GAMs are computed (Hastie and Tibshirani 1990), hence justifying

the use of AIC in the context of GAMs.

Chapter 3

Incorporating covariates into standard line transect analysis

3.1 Introduction

Standard line transect data consist of records of the amount of effort associated with each transect surveyed, as well as the number of sightings, perpendicular distance of each sighting to the transect line, and the cluster size of each sighting (if applicable). In addition, auxiliary (covariate) information is also routinely recorded. Such covariate data can be classified into three categories: those associated with detections only (cluster size, for example, can be thought of as a covariate); those associated with the effort (e.g. Beaufort, search equipment used); and spatial covariates that are available for the whole study region (e.g. altitude/depth, oceanographic variables). The latter are usually obtained independently from the line transect survey. Covariate information can potentially be incorporated into the estimation of encounter rate, mean school size, or $f(0)$. In Chapter 2 we presented a method that uses spatial covariates to model encounter rate and mean school size in space and time. In this Chapter we address the issue of incorporating covariate information into the estimation of $f(0)$ which, as with the spatio-temporal modeling approach described earlier, may be particularly useful for the case where data were obtained from opportunistic surveys.

Standard line transect methodology assumes that the probability of detection of an object depends solely on its perpendicular distance from the transect line. In practice, it is widely known that a number of variables may affect the detection probability (e.g. Caughley 1974, Gunnlaugsson and Sigurjónsson 1990). Attempts to minimise or eliminate heterogeneity in detection probabilities are usually carried out in two ways (Buckland *et al.* 1993). One

way is by stratifying the data according to some set of covariate values and separately estimating $f(0)$ for each stratum. Stratification can be carried out before the survey takes place, when *a priori* information is available, or can be based on examination of the data after the survey has been conducted (post-stratification). When object density is low, however, small sample sizes may preclude stratification or result in biased estimates of $f(0)$, and hence of density, if stratification is carried out. A second approach is to incorporate covariates into $f(0)$ estimation.

Beavers and Ramsey (1998) proposed regressing the logarithm of the observed perpendicular distances on the covariates and adjusting these distances to average covariate values using the estimated regression parameters. The main advantage of this approach is that it allows standard software to be used for the estimation of $f(0)$, by analysing the adjusted distances (e.g. Fancy 1997). However, it is somewhat *ad hoc* in two respects. Firstly, it requires some arbitrary small value to be added to perpendicular distances equal to zero. Secondly, it relies on a common truncation distance to be applied to all adjusted distances, irrespective of their associated covariate values. This may introduce bias in the resulting estimate of $f(0)$, as it is unlikely that the tail of the distribution of the adjusted observed perpendicular distances will encompass similar proportions of the various covariate values associated with the detected objects. Thus truncation of those extreme adjusted distances at the tail of the distribution may result in a systematically higher proportion of observations having a few specific covariate values being excluded from the analyses.

Alternatively, the covariates can be directly incorporated into the estimation procedure via a multivariate detection function. This requires estimation of the joint density of the observed perpendicular distances x and associated covariates $\mathbf{z}$ ($\mathbf{z} = z_1, \ldots, z_q$) (Appendix 3.7.1):

$$f(x, \mathbf{z}) = \frac{g(x, \mathbf{z})\, \pi(x, \mathbf{z})}{\int_{\mathbf{Z}} \int_X g(x, \mathbf{z})\, \pi(x, \mathbf{z})\, dx\, d\mathbf{z}}, \tag{3.1}$$

where $g(x, \mathbf{z})$ is a multivariate detection function, $\pi(x, \mathbf{z})$ is the joint density of x and $\mathbf{z}$ in the population, and integration is done over the range of the indicator variable. Assuming that the x and $\mathbf{z}$ are independent, then $\pi(x, \mathbf{z}) = \pi(x)\, \pi(\mathbf{z})$, where $\pi(x)$ and $\pi(\mathbf{z})$ denote the densities of the x's and $\mathbf{z}$'s, respectively, so that we have:

$$f(x, \mathbf{z}) = \frac{g(x, \mathbf{z})\, \pi(x)\, \pi(\mathbf{z})}{\int_{\mathbf{Z}} \int_X g(x, \mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, d\mathbf{z}}. \tag{3.2}$$

Given random line placement, $\pi(x) = 1/W$, $0 \leq x \leq W$, and the expression above is reduced to:

$$f(x, \mathbf{z}) = \frac{g(x, \mathbf{z})\, \pi(\mathbf{z})}{\int_{\mathbf{Z}} \pi(\mathbf{z}) \int_X g(x, \mathbf{z})\, dx\, d\mathbf{z}}. \tag{3.3}$$

$\pi(\mathbf{z})$ is usually not known, and so must be either estimated or factored out. For univariate z, Chen (1996) proposed a bivariate density estimator based on the product of two Gaussian kernels. The density $f(x, z)$ is then directly estimated from the data, without the need to assume any parametric form. A similar method based on a single Gaussian kernel has been proposed by Mack and Quang (1998). Both methods were developed with the primary aim of estimating mean school size and its effect on detectability. Although they seem to perform relatively well, they suffer from a few disadvantages. One difficulty is the choice of bandwidth h. For the estimators of Chen and Mack and Quang, $h = \hat{\sigma} n^{-\delta}$, where $\hat{\sigma}$ is the sample estimate of the standard error, n corresponds to the total number of detected objects, and δ is a constant. The choice of value for δ depends on the criterion being used to derive unbiased distributional properties. As the rate of shrinkage of the bandwidth varies according to the criterion being used, confidence intervals may be biased (Mack and Quang 1998). In addition, density estimates may depend on the choice of bandwidth (e.g. Buckland 1992*a*). Secondly, as the kernel method is based on local averaging of observations, it is more likely to produce biased estimates when the detection function is not very smooth near $x = 0$ (see, for example, Buckland 1992*a*). Finally, at least one of the methods (Mack and Quang 1998) requires relatively large sample sizes ($n \geq 70$) for unbiased estimation of mean school size.

Alternatively, maximum likelihood estimation of the conditional density of the x's given the observed $\mathbf{z}$'s also does not require knowledge of $\pi(\mathbf{z})$. Following the derivation of Borchers (1996) we have:

$$f(\mathbf{z}) = \int_X f(x, \mathbf{z})\, dx = \frac{\pi(\mathbf{z}) \int_X g(x, \mathbf{z})\ \pi(x)\ dx}{\int_{\mathbf{Z}} \int_X g(x, \mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, d\mathbf{z}} \tag{3.4}$$

and, given random line placement so that $\pi(x) = 1/W$, then

$$f(x \mid \mathbf{z}) = \frac{f(x, \mathbf{z})}{f(\mathbf{z})} = \frac{g(x, \mathbf{z})}{\int_X g(x, \mathbf{z})\ dx}, \tag{3.5}$$

so that the conditional likelihood is given by:

$$\mathcal{L}(\theta; \mathbf{x}, \mathbf{z}) = \prod_{i=1}^{n} f(x_i \mid \mathbf{z}_i) = \prod_{i=1}^{n} \frac{g(x_i, \mathbf{z}_i)}{\int_X g(x, \mathbf{z}_i)\, dx}. \tag{3.6}$$

A number of authors have used this conditional approach to develop estimators for $f(0)$ (Drummer and McDonald 1987, Ramsey *et al.* 1987, Quang 1991). Quang (1991) modified the Fourier series model of Crain *et al.* (1979) to include an additional variable other than perpendicular distance. All other estimators assume that the covariates enter the detection function via the scale parameter. Empirical evidence (Otto and Pollock 1990) supports this contention, as long as detection on the line is certain (*i.e.* $g(0, \mathbf{z}) = 1$). In the case where covariates affect detection on the line, methods that estimate $g(0, \mathbf{z})$ (e.g. Borchers *et al.* 1998*a*) should be used instead.

Ramsey *et al.* (1987) formulated an estimator for variable-area circular plot surveys based on the density of detected areas y, where $y = \pi x^2$ in the case of point transects. In the context of line transect sampling, modelling of detected areas is equivalent to the modelling of the observed perpendicular distances themselves. Thus, following the notation used throughout this Chapter, and using the observed perpendicular distances in place of detected areas, Ramsey *et al.*'s estimator has the form $g(x, \mathbf{z}) = h(x/\mu(\mathbf{z}))$. Here $\mu(\mathbf{z}) = \int_X g(x, \mathbf{z})\, dx$ and $h(\cdot)$ is a detection function with properties such that the conditional density of the observed perpendicular distances given the associated covariates, $\tau(x \mid \mathbf{z})$, is given by:

$$\tau(x \mid \mathbf{z}) = \frac{h(x/\mu(\mathbf{z}))}{\mu(\mathbf{z})}. \tag{3.7}$$

Thus Ramsey *et al.*'s estimator is equivalent to the conditional expression from (3.5). By assuming that $\log_e(\mu(\mathbf{z})) = \beta_0 + \sum_{j=1}^{q} \beta_j z_j$, and that $\tau(x \mid \mathbf{z})$ follows a Weibull distribution with a single shape parameter γ, $\tau(x \mid \mathbf{z})$ is then estimated by maximising the log-likelihood with respect to the $q+1$ covariate parameters and the shape parameter γ, conditional on the observed values of the covariates.

The bivariate density estimator proposed by Drummer and McDonald (1987) is based on a family of models which satisfy the assumption that $\int_X g(x, z)\, dx = c \cdot \int_X g(x)\, dx$, with $c = z^{\beta}$. As in Ramsey *et al.*'s formulation, this approach is a special case of expression (3.5).

It can easily be seen that both $\mu(\mathbf{z})$ and c, denoted in the expression below by σ, have the form (Borchers 1996):

$$\sigma = exp\left(\beta_0 + \sum_{j=1}^{q} \beta_j t(z_j)\right), \tag{3.8}$$

with $t(z_j) = z_j$ for Ramsey *et al.*'s $\mu(\mathbf{z})$, and $\beta_0 = 0$, $t(z_j) = \log_e(z_j)$ and $q = 1$ for Drummer and McDonald's c. Palka (1993) applied Drummer and McDonald's method parameterised as in (3.8) to fit a bivariate detection function based on the hazard-rate model of Hayes and Buckland (1983) and estimate the effect of school size on the detectability of harbour porpoise.

The conditional likelihood methods described above are based on either a series representation (Quang 1991) or a specific parametric functional form (Drummer and McDonald 1987, Ramsey *et al.* 1987) for the detection function $g(x, z)$ (Borchers 1996). Buckland (1992*a*,*b*) developed a unifying framework where the two approaches are combined to model $f(0)$ in the case where the probability of detection is a function of the perpendicular distances alone. Buckland's approach involves the fitting of a known parametric form (the 'key function') to the data, with additional adjustment terms used when necessary to improve the fit. This semiparametric approach is very flexible, and so is 'model robust' in the sense of Burnham *et al.* (1980). In addition, as model parameters are estimated by maximum likelihood, objective model selection criteria can be employed (Borchers 1996).

We propose a generalisation of Buckland's approach in which covariates are incorporated into the estimation of the detection probabilities via the scale parameter σ. In this formulation, the covariates are assumed to affect the rate at which detectability decreases as a function of distance, but they do not influence the shape of the detection function. Unlike kernel-based methods, this approach does not involve local fitting, and so it is more robust to rounding of measurements. Also, confidence interval estimation is not sensitive to the parameterisation used. Finally, the method allows for the inclusion of more than one covariate (other than perpendicular distance), and it is easy to implement.

3.2 A unifying framework for incorporating covariates into the estimation of detection probabilities

Let there be a set of transect lines of total length L placed over a survey region of area A according to some survey design. An observer travels along each transect and records the perpendicular distance x_i and covariate values $\mathbf{z}_i$ ($\mathbf{z}_i = z_{1i}, \ldots, z_{qi}$) associated with each detected object i ($i = 1, \ldots, n$). Only objects located up to distance W from the line are recorded. The usual assumptions of line transect methodology are deemed to hold (see Chapter 1 and Buckland *et al.* 1993 for a review of assumptions and implications of violations). Of primary importance are the assumptions that (i) all objects located on or near the line are detected with certainty (i.e. $g(0, \mathbf{z}) = 1$); (ii) objects are detected prior to any responsive movement; and (iii) measurements are made without errors.

Following the 'key function' formulation of Buckland (1992*a*), and using the result from (3.5), the conditional density $f(x \mid \mathbf{z})$ is given by:

$$\begin{aligned} f(x \mid \mathbf{z}) &= \frac{k(x, \mathbf{z})}{\int_0^W k(x, \mathbf{z}) \left[1 + \sum_{j'=1}^m a_{j'} p_{j'}(x_s)\right] dx} \left[1 + \sum_{j'=1}^m a_{j'} p_{j'}(x_s)\right] \\ &= \frac{k(x, \mathbf{z})}{\mu(\mathbf{z})} \left[1 + \sum_{j'=1}^m a_{j'} p_{j'}(x_s)\right]. \end{aligned} \tag{3.9}$$

Here $k(x, \mathbf{z})$ is a parametric function (e.g. half-normal or hazard-rate), $p_{j'}$ is an adjustment term (cosine, simple or Hermite polynomials) of order j' ($j' = 1, \ldots, m$), $a_{j'}$ is the coefficient for the j'th adjustment term, and x_s is a standardised x value required to avoid numerical difficulties (Buckland 1992*a*) and taken to be x/σ, where σ is the scale term (see below). The integral $\int_0^W k(x, \mathbf{z}) \left[1 + \sum_{j'=1}^m a_{j'} p_{j'}(x_s)\right] dx$, in which W denotes the truncation distance, is a normalising function of $\mathbf{z}$ and the parameters, required to ensure that $f(x \mid \mathbf{z})$ integrates to 1. We consider just the half-normal and hazard-rate key functions; other key functions available in DISTANCE (Laake *et al.* 1993) either do not allow the inclusion of covariates (uniform key) or have an implausible shape close to $x = 0$ (exponential key), and hence are not considered here.

Assume that the covariates affect detectability via the scale term σ. We investigate two functional forms for σ, namely:

$$\sigma_i = exp\left(\beta_0 + \sum_{j=1}^{q} \beta_j z_{ij}\right), \tag{3.10}$$

where the covariate effects are multiplicative as previously assumed by other authors (*cf.* expression (3.8)), and a linear form:

$$\sigma_i = \beta_0 + \sum_{j=1}^{q} \beta_j z_{ij}. \tag{3.11}$$

Parameter estimates are obtained by maximising the conditional log-likelihood

$$\ell = \log_e [\mathcal{L}(\theta; \mathbf{x}, \mathbf{z})] = \log_e \left[\prod_{i=1}^{n} f(x_i \mid \mathbf{z}_i)\right] = \sum_{i=1}^{n} \log_e [f(x_i \mid \mathbf{z}_i)] \tag{3.12}$$

with respect to the parameter vector θ, where $\theta = \theta_1, \ldots, \theta_{k+q+1+m}$, and k, $q+1$ and m refer to the number of shape parameters of the key function, scale and adjustment terms, respectively. Note that $k = 0$ for the half-normal key, whereas for the hazard-rate key function $k = 1$ (the 'power' parameter). Defining

$$t(x, \mathbf{z}) = k(x, \mathbf{z}) \left[1 + \sum_{j'=1}^{m} a_{j'} p_{j'}(x_s)\right], \tag{3.13}$$

we have

$$\begin{aligned} \ell &= \sum_{i=1}^{n} \log_e \left[\frac{t(x_i, \mathbf{z}_i)}{\mu(\mathbf{z}_i)}\right] \\ &= \sum_{i=1}^{n} \log_e [t(x_i, \mathbf{z}_i)] - \sum_{i=1}^{n} \log_e [\mu(\mathbf{z}_i)] \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} \frac{\partial \ell}{\partial \theta_j} &= \sum_{i=1}^{n} \frac{\partial \log_e [t(x_i, \mathbf{z}_i)]}{\partial \theta_j} - \sum_{i=1}^{n} \frac{\partial \log_e [\mu(\mathbf{z}_i)]}{\partial \theta_j} \\ &= \sum_{i=1}^{n} \frac{1}{t(x_i, \mathbf{z}_i)} \frac{\partial t(x_i, \mathbf{z}_i)}{\partial \theta_j} - \sum_{i=1}^{n} \frac{1}{\mu(\mathbf{z}_i)} \frac{\partial \mu(\mathbf{z}_i)}{\partial \theta_j}. \end{aligned} \tag{3.15}$$

For $\theta_j = \theta_{k+1}, \ldots, \theta_{k+q+1}$ (*i.e.* the parameters of the scale term), it is convenient to rewrite the equation above as:

$$\frac{\partial \ell}{\partial \theta_j} = \sum_{i=1}^{n} \frac{1}{t(x_i, \mathbf{z}_i)} \frac{\partial t(x_i, \mathbf{z}_i)}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \theta_j} - \sum_{i=1}^{n} \frac{1}{\mu(\mathbf{z}_i)} \frac{\partial \mu(\mathbf{z}_i)}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \theta_j} \tag{3.16}$$

so that now the only additional computation required in comparison with the standard estimation approach are the partial derivatives of the σ_i with respect to θ_j (*i.e.* the β_js from expressions (3.10) and (3.11)).

We use the algorithm described by Buckland (1992*a*,*b*), extended to include the covariate parameters, to estimate the model parameters. It uses the Newton-Raphson method with a Marquardt procedure to fit the key function to the data. Once the key function has been fitted, polynomial terms are fitted conditional on the estimated key parameters. Another iteration is then carried out to estimate the key function parameters, conditioning on the estimated polynomial coefficients. This procedure is repeated until it approaches convergence, at which stage all parameters are maximised simultaneously. We use initial parameter estimates as described in Buckland (1992*b*), with the sample variance and the logarithm of the sample variance being used as initial estimates of β_0 under the linear and exponential formulations for the scale term, respectively, and all other coefficients corresponding to the covariate parameters set equal to zero.

Selection of the number of adjustment terms to be included in the model can be carried out based on either likelihood ratio tests or Akaike's information criterion (AIC; Akaike 1973). Variance estimates are obtained from the Hessian matrix using the final parameter estimates (Buckland 1992*a*).

The method described above is specific to ungrouped line transect data, although the approach can be easily extended to point transects and to grouped (line transect or point transect) data (Appendix 3.7.2).

3.3 Abundance estimation

3.3.1 Single objects

The standard univariate (*i.e.* based on perpendicular distances alone) line transect estimator of abundance is given by (Buckland *et al.* 1993):

$$\hat{N}_s = A \cdot \hat{D}_s = \frac{A}{2L} \cdot n \cdot \hat{f}(0), \tag{3.17}$$

where $\hat{N}_s$ is an estimate of the total abundance N_s, $\hat{D}_s$ is an estimate of the density of objects D_s, and n is the total number of detections (see Chapter 1, Seber (1982) and Buckland *et al.* (1993) for the derivation of the above estimator). The variance of $\hat{N}_s$ is obtained using the delta method (Seber 1982, p.7), and it can be estimated by (Buckland *et al.* 1993):

$$\begin{aligned} \widehat{var}(\hat{N}_s) &= A^2 \widehat{var}(\hat{D}_s) \\ &= A^2 \cdot \hat{D}_s^2 \left\{ \frac{\widehat{var}\left\{\hat{f}(0)\right\}}{\left\{\hat{f}(0)\right\}^2} + \frac{\widehat{var}(n)}{n^2} \right\}. \end{aligned} \tag{3.18}$$

To derive an estimator for the case where we have a multivariate conditional density $f(0 \mid \mathbf{z})$, it is convenient to view the above abundance estimator as a Horvitz-Thompson (Horvitz and Thompson 1952) like estimator. Horvitz-Thompson estimators are based on inclusion probabilities, and do not require the assumption that all objects have the same probability of being included in the sample. Hence, in the context of line transect sampling, define p_i to be the probability that object i is detected within the strip of width W, conditional on the observed values of $\mathbf{z}_i$; that is:

$$\begin{aligned} p_i &= \mathbb{E}\left[g(x_i, \mathbf{z}_i) \mid \mathbf{z}_i\right] \\ &= \int_0^W g(x, \mathbf{z}_i)\, \pi(x \mid \mathbf{z}_i)\, dx, \end{aligned} \tag{3.19}$$

where $g(x, \mathbf{z}_i) = f(x \mid \mathbf{z}_i)/f(0 \mid \mathbf{z}_i)$. Assuming the x and $\mathbf{z}$ are independent, and that objects are uniformly distributed relative to the transect line, then $\pi(x \mid \mathbf{z}_i) = \pi(x) = 1/W$. Thus:

$$
\begin{aligned}
p_i &= \int_0^W g(x, \mathbf{z}_i)\, \pi(x)\, dx \\
&= \frac{1}{W} \int_0^W g(x, \mathbf{z}_i)\, dx \\
&= \frac{1}{W} \frac{1}{f(0 \mid \mathbf{z}_i)}.
\end{aligned}
\tag{3.20}
$$

Taking the p_i as known, a Horvitz-Thompson estimator of M_s, the total number of objects within the area surveyed, is then given by:

$$
\hat{M}_s = \sum_{i=1}^{n} \frac{1}{p_i} = W \sum_{i=1}^{n} f(0 \mid \mathbf{z}_i), \tag{3.21}
$$

with a Horvitz-Thompson-like estimator of the total abundance N_s in the survey region given by (Borchers 1996, Borchers *et al.* 1998*b*):

$$
\begin{aligned}
\hat{N}_{HTs} &= \frac{A}{2LW} \hat{M}_s \\
&= \frac{A}{2L} \sum_{i=1}^{n} f(0 \mid \mathbf{z}_i).
\end{aligned}
\tag{3.22}
$$

Taking the expectation of $f(0 \mid \mathbf{z})$ with respect to the distribution of the observed $\mathbf{z}$ we have (Borchers 1996):

$$
\mathbb{E}_{\mathbf{z}} [f(0 \mid \mathbf{z})] = \int_{\mathbf{Z}} \frac{g(0, \mathbf{z})}{\int_0^W g(x, \mathbf{z})\, dx} f(\mathbf{z})\, d\mathbf{z}. \tag{3.23}
$$

Using the definition of $f(\mathbf{z})$ from expression (3.4), and remembering that $\int_{\mathbf{Z}} \pi(\mathbf{z}) = 1$, and that by assumption $g(0, \mathbf{z})$ must equal 1, we then have:

$$
\mathbb{E}_{\mathbf{z}} [f(0 \mid \mathbf{z})] = \frac{\pi(x)}{\int_{\mathbf{Z}} \int_X g(x, \mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, d\mathbf{z}}. \tag{3.24}
$$

Given random line placement $\pi(x) = 1/W$ and so we have:

$$\mathbb{E}_{\mathbf{z}}\left[f(0 \mid \mathbf{z})\right] = \frac{1}{W \int_{\mathbf{Z}} \int_{X} g(x,\mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, dz}. \tag{3.25}$$

Recall that:

$$\begin{aligned} f(0) &= \int_{\mathbf{Z}} f(0,\mathbf{z})\, dz \\ &= \int_{\mathbf{Z}} \frac{g(0,\mathbf{z})\, \pi(0)\, \pi(\mathbf{z})\, dz}{\int_{\mathbf{Z}} \int_{X} g(x,\mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, dz} \end{aligned} \tag{3.26}$$

and, since by assumption $g(0,\mathbf{z}) = 1$, then:

$$f(0) = \frac{\pi(0)}{\int_{\mathbf{Z}} \int_{X} g(x,\mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, dz} \int_{\mathbf{Z}} \pi(\mathbf{z})\, dz. \tag{3.27}$$

But since $\pi(0) = \pi(x) = 1/W$, then:

$$f(0) = \frac{1}{W \int_{\mathbf{Z}} \int_{X} g(x,\mathbf{z})\, \pi(x)\, \pi(\mathbf{z})\, dx\, dz}. \tag{3.28}$$

Therefore the result from expression (3.25) for the conditional expectation of $f(0)$ given the covariates, is $f(0)$. Thus, given an estimate of $f(0 \mid \mathbf{z})$, we can then obtain an estimate of $f(0)$ as the average of the $\hat{f}(0 \mid \mathbf{z}_i)$ (Borchers 1996).

Horvitz-Thompson estimators are unbiased (Thompson 1992). Thus, conditional on n, replacing $f(0 \mid \mathbf{z})$ by its estimator $\hat{f}(0 \mid \mathbf{z})$ in expression (3.22) yields an unbiased estimate of N_s, as long as the estimates of $f(0 \mid \mathbf{z}_i)$ are unbiased (Borchers 1996).

Under the assumption that detections are independent, the variance for $\hat{M}_s$ is given by:

$$var(\hat{M}_s) = \sum_{i=1}^{n} \frac{1 - p_i}{p_i}, \tag{3.29}$$

which can be estimated by (Thompson 1992):

$$\widehat{var}(\hat{M}_s) = \sum_{i=1}^{n} \frac{1 - p_i}{p_i^2}. \tag{3.30}$$

Substituting p_i by its estimator yields the follow expression for the estimate of the variance of $\hat{M}_s$ (Borchers 1996):

$$\widehat{var}(\hat{M}_s) = W^2 \sum_{i=1}^{n} \hat{f}(0 \mid \mathbf{z}_i)^2 - \hat{M}_s. \tag{3.31}$$

As the estimator $\hat{N}_{HTs}$ is conditional on the observed covariate values and parameter estimates, using standard results for conditional variances (Seber 1982, p.9) we have:

$$var(\hat{N}_{HTs}) = \mathbb{E}_{\hat{\theta}}\left[var\left(N_{HTs} \mid \hat{\theta}\right)\right] + var_{\hat{\theta}}\left(\mathbb{E}\left[N_{HTs} \mid \hat{\theta}\right]\right) \tag{3.32}$$

which can be estimated by:

$$\widehat{var}(\hat{N}_{HTs}) = \left(\frac{A}{2LW}\right)^2 \left\{ W^2 \sum_{i=1}^{n} \hat{f}(0 \mid \mathbf{z}_i)^2 - \hat{M}_s + \sum_{j=1}^{r}\sum_{l=1}^{r} \frac{\partial \hat{M}_s}{\partial \hat{\theta}_j}^T H(\hat{\theta}_{jl})^{-1} \frac{\partial \hat{M}_s}{\partial \hat{\theta}_l} \right\} \tag{3.33}$$

where r in the summations refers to all $k+q+1+m$ parameters and $H(\hat{\theta}_{jl})^{-1}$ denotes the jlth element of the inverse of the Hessian matrix, which is given by (Buckland *et al.* 1993):

$$H(\hat{\theta}_{jl}) = \frac{1}{n}\left[\sum_{i=1}^{n} \frac{\partial \log_e[f(0 \mid \mathbf{z}_i)]}{\partial \hat{\theta}_j} \cdot \frac{\partial \log_e[f(0 \mid \mathbf{z}_i)]}{\partial \hat{\theta}_l}\right]. \tag{3.34}$$

An advantage of viewing the estimator of N_s as a Horvitz-Thompson-like estimator is that its estimate of the variance, given by the expression above, incorporates the variance component due to estimation of the parameters of the detection function. This avoids the negative bias common to Drummer and McDonald's (1987) and Quang's (1991) estimators of precision (Borchers 1996), which effectively treat the estimated parameters as known. Alternatively, the variance of $\hat{N}_{HTs}$ can be estimated using the bootstrap (Efron and Tibshirani 1993). Assuming that detections from different transect lines are independent, the transects can be taken to be the sampling units, and the procedure would then be as follows: At each of B bootstrap iterations, resample transect lines along with their corresponding detections, with replacement, until the total amount of effort from the resampled lines approximates the original total effort. Estimates of $f(0 \mid \mathbf{z})$ can then

be obtained using the methodology described in this section, and abundance estimates obtained using expression (3.22). If at each iteration model selection is carried out, then this approach has the advantage of incorporating model selection uncertainty into the estimate of precision.

3.3.2 Objects in clusters

In the case where objects occur in clusters, the standard line transect estimator of abundance is given by (Buckland *et al.* 1993):

$$\hat{N} = A \cdot \hat{D} = \frac{A}{2L} \cdot n \cdot \hat{\mathbb{E}}[s] \cdot \hat{f}(0), \tag{3.35}$$

where $\hat{\mathbb{E}}[s]$ denotes the estimated mean cluster size of the population. An estimator of the variance of $\hat{N}$ is then given by (Buckland *et al.* 1993):

$$\begin{aligned} \widehat{var}(\hat{N}) &= A^2 \widehat{var}(\hat{D}) \\ &= A^2 \cdot \hat{D}^2 \left\{ \frac{\widehat{var}\left\{\hat{f}(0)\right\}}{\left\{\hat{f}(0)\right\}^2} + \frac{\widehat{var}(\hat{\mathbb{E}}[s])}{\left\{\hat{\mathbb{E}}[s]\right\}^2} + \frac{\widehat{var}(n)}{n^2} \right\}. \end{aligned} \tag{3.36}$$

A Horvitz-Thompson estimator of the total number of clusters within the area surveyed is given by expression (3.21). A Horvitz-Thompson estimator of the total number of objects within the area surveyed, M, is given by (Thompson 1992):

$$\hat{M} = \sum_{i=1}^{n} \frac{s_i}{p_i} \tag{3.37}$$

where s_i denotes the size of the ith detected cluster. Following the derivation from the previous section we then have:

$$\hat{M} = W \sum_{i=1}^{n} s_i \cdot f(0 \mid \mathbf{z}_i). \tag{3.38}$$

An estimate of the overall abundance of objects in the survey region, $\hat{N}_{HT}$, is then obtained by substituting the expression above into equation (3.22), so that:

$$\begin{aligned}\hat{N}_{HT} &= \frac{A}{2LW}\hat{M} \\ &= \frac{A}{2L}\sum_{i=1}^{n} s_i \cdot f(0 \mid \mathbf{z}_i).\end{aligned} \tag{3.39}$$

As in the case of single objects, substituting $f(0 \mid \mathbf{z}_i)$ by its estimator $\hat{f}(0 \mid \mathbf{z}_i)$ will yield unbiased estimates of abundance as long as the $\hat{f}(0 \mid \mathbf{z}_i)$ are unbiased.

Assuming that detections are independent, an estimate of the variance of $\hat{M}$ is given by (Thompson 1992):

$$\begin{aligned}\widehat{var}(\hat{M}) &= \sum_{i=1}^{n}\left(\frac{1-p_i}{p_i^2}\right)s_i^2 \\ &= \sum_{i=1}^{n}\left\{W^2\hat{f}(0 \mid \mathbf{z}_i)^2 - W\hat{f}(0 \mid \mathbf{z}_i)\right\}s_i^2,\end{aligned} \tag{3.40}$$

with an estimate of the variance of the total abundance of objects in the survey region given by:

$$\begin{aligned}\widehat{var}(\hat{N}_{HT}) &= \left(\frac{A}{2LW}\right)^2\left\{\sum_{i=1}^{n}\left\{W^2\hat{f}(0 \mid \mathbf{z}_i)^2 - W\hat{f}(0 \mid \mathbf{z}_i)\right\}s_i^2\right. \\ &\qquad \left. + \sum_{j=1}^{r}\sum_{l=1}^{r}\frac{\partial\hat{M}^T}{\partial\hat{\theta}_j}H(\hat{\theta}_{jl})^{-1}\frac{\partial\hat{M}}{\partial\hat{\theta}_l}\right\}.\end{aligned} \tag{3.41}$$

Note that an estimate of the expected mean cluster size $\mathbb{E}[s]$ is readily obtained as:

$$\begin{aligned}\hat{\mathbb{E}}[s] &= \frac{\hat{M}}{\hat{M}_s} \\ &= \frac{\sum_{i=1}^{n} s_i \cdot \hat{f}(0 \mid \mathbf{z}_i)}{\sum_{i=1}^{n}\hat{f}(0 \mid \mathbf{z}_i)},\end{aligned} \tag{3.42}$$

with an estimate of its variance obtained using the delta method, so that:

$$\widehat{var}(\hat{\mathbb{E}}[s]) = \hat{\mathbb{E}}[s]^2 \left\{ \frac{\widehat{var}(\hat{M}_s)}{\hat{M}_s^2} + \frac{\widehat{var}(\hat{M})}{\hat{M}^2} - \frac{2\,\widehat{cov}(\hat{M}_s, \hat{M})}{\hat{M}_s \hat{M}} \right\}, \tag{3.43}$$

where $\widehat{var}(\hat{M}_s)$ and $\widehat{var}(\hat{M})$ are obtained from expressions (3.31) and (3.40), and $\widehat{cov}(\hat{M}_s, \hat{M})$ can be approximated by (Seber 1982, p.9):

$$\widehat{cov}(\hat{M}_s, \hat{M}) = \sum_{l=1}^{q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \widehat{cov}(z_{li}, z_{lj}) \frac{\partial \hat{M}_s}{\partial z_{li}} \frac{\partial \hat{M}}{\partial z_{lj}}.$$

However, estimation of $cov(z_{li}, z_{lj})$ requires knowledge of the pdf of the covariates $\mathbf{z}$. Alternatively the variance of $\hat{\mathbb{E}}[s]$ can be obtained by bootstrapping the transect lines as described at the end of section 3.3.1. At each bootstrap iteration, $\hat{M}_s$ and $\hat{M}$ can be computed based on the resampled data, and an estimate of the mean cluster size can be obtained using equation (3.42). Under this approach it is assumed that transect lines are independent, but detected objects and their cluster size need not be. Hence we can directly estimate the variance of $\hat{\mathbb{E}}[s]$ from its bootstrap estimates.

3.4 Simulations

3.4.1 Size-bias

The most common scenario addressed in the literature is when the cluster size associated with a particular object influences the probability of that object being detected, which is commonly referred to as size-bias. Size-bias is of special concern since it may result in a biased estimate of the expected cluster size, which in turn may lead to a biased estimate of the density of objects (see pp. 125-135 in Buckland *et al.* (1993) for a detailed discussion).

To investigate the statistical properties of the estimator derived in section 3.2, and to allow a comparison with the performance of alternative estimators developed explicitly to deal with size-bias, we follow the simulation procedure of Chen (1996). We simulate a population N of 300 objects uniformly distributed in the interval $[0, W]$, with $W = 10$. To ensure that the density of objects in the area surveyed $D = 1.0$, we specify the survey area to be equal to $2LW$, with the transect length $L = 15$. The cluster size s associated with each object is generated from a χ^2_5 distribution. Size-bias is introduced via the bivariate detection function of Drummer and McDonald (1987), namely:

$$g(x,s) = exp\left(-b\left(\frac{x}{s^c}\right)^a\right), \tag{3.44}$$

which is used to obtain the detection probabilities for each simulated object. The rejection method (*cf.* Ripley 1987) is then applied to determine which objects are detected. As in Chen (1996), we fix $b = 0.5$ and vary the parameters a and c (*cf.* Table 3.1). The density of objects D within the area surveyed is estimated as $\hat{D} = n \cdot \hat{f}(0)/2L$, where $\hat{f}(0) = \sum_{i=1}^{n} \hat{f}(0 \mid \mathbf{z}_i)/n$. An estimate of the mean cluster size $\hat{\mathbb{E}}[s]$ is obtained according to expression (3.42). A total of 240 simulations, with 119 bootstrap iterations carried out for each simulation, were performed. Estimates of bias, standard errors, root mean squared errors (RMSE) and coverage of the density and mean school size estimators were computed using the simulation estimates. Choice of the number of simulations and bootstrap iterations to be performed was based on the need to reduce simulation error relative to statistical error (*cf.* Davison and Hinkley 1997), and also taking into account computational costs. We used the model fitted to the original data in all bootstrap iterations and, as Chen (1996), condition the bootstrap estimates on the simulated number of detected objects (n). Bootstrap iterations in which no convergence was achieved were discarded and new iterations carried out to ensure a total of 119 iterations for each simulation. Independent simulations were carried out to evaluate the estimators based on the scale term as an exponential function (expression (3.10)) and as a linear function (expression (3.11)). The latter did not perform well in comparison with the former, and so we only present simulation results for the estimator based on an exponential form for the scale parameter (Table 3.1).

Point estimates of density of clusters ($\hat{D}$) and mean cluster size ($\hat{\mathbb{E}}[s]$) were close to their true values, although they both showed a slight upward bias. Estimates of standard error varied little between the various combinations of parameter values, with the exception of the estimate for simulated data with a small shoulder (a = 1.5) and small size-bias effect (c = 0.2), which presented an estimate of standard error more than twice as large as that of all others. Examination of simulated estimates for this combination of parameter values revealed a single estimate of D ($\hat{D} = 5.34$) and $\mathbb{E}[s]$ ($\hat{\mathbb{E}}[s] = 15.53$) much greater than the observed average. Omitting this outlier results in estimates of bias and standard error closer to the observed values for the other parameter cases ($\hat{D} = 1.02$, RMSE($\hat{D}$) = 0.1011, $\widehat{Bias}(\hat{D}) = 0.0157$, $\widehat{SE}(\hat{D}) = 0.0988$, $\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D}) = 0.1589$, $\hat{\mathbb{E}}[s] = 5.10$, RMSE($\hat{\mathbb{E}}[s]$) = 0.4261, $\widehat{Bias}(\hat{\mathbb{E}}[s]) = 0.0969$, $\widehat{SE}(\hat{\mathbb{E}}[s]) = 0.4043$, $\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s]) = 0.2396$). For both estimators coverage was satisfactory, generally near the nominal value

	$a = 1.5$		$a = 2.0$		$a = 2.5$	
	$c = 0.2$	$c = 0.6$	$c = 0.2$	$c = 0.6$	$c = 0.2$	$c = 0.6$
$\hat{D}$	1.0337	1.0251	1.0252	1.0436	1.0262	1.0481
RMSE($\hat{D}$)	0.2992	0.0859	0.0912	0.1006	0.0826	0.1198
$\widehat{Bias}(\hat{D})$	0.0337	0.0251	0.0252	0.0436	0.0262	0.0481
$\widehat{SE}(\hat{D})$	0.2960	0.0784	0.0841	0.0796	0.0740	0.0988
$\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$	0.1139	0.3205	0.3002	0.5476	0.3538	0.4868
Coverage($\hat{D}$)	0.9542	0.9583	0.9417	0.9458	0.9583	0.8792
$\hat{\mathbb{E}}[s]$	5.1403	5.0860	5.1113	5.0998	5.0574	5.1199
RMSE($\hat{\mathbb{E}}[s]$)	0.8084	0.2916	0.3429	0.2695	0.2979	0.4643
$\widehat{Bias}(\hat{\mathbb{E}}[s])$	0.1403	0.0860	0.1113	0.0998	0.0574	0.1199
$\widehat{SE}(\hat{\mathbb{E}}[s])$	0.7852	0.2656	0.3053	0.2300	0.2872	0.4331
$\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$	0.3279	0.3498	0.3080	0.3953	0.2606	0.3388
Coverage($\hat{\mathbb{E}}[s]$)	0.9583	0.9625	0.9583	0.9542	0.9417	0.9500

Table 3.1: Estimates of density of clusters ($\hat{D}$) and mean cluster size ($\hat{\mathbb{E}}[s]$) and their corresponding estimates of bias ($\widehat{Bias}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])$), standard error ($\widehat{SE}(\hat{D})$ and $\widehat{SE}(\hat{\mathbb{E}}[s])$), root mean squared error (RMSE ($\hat{D}$) and RMSE ($\hat{\mathbb{E}}[s]$)), the ratio of the estimate of bias to the estimate of standard error ($\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$), and coverage (Coverage($\hat{D}$) and Coverage($\hat{\mathbb{E}}[s]$)) obtained using the method described in section 3.2, with the scale parameter as an exponential function, applied to simulated data. Here 'a' indicates values for the shape ('power') parameter, whereas 'c' corresponds to parameter values of the scale term. Note that RMSE $= \sqrt{\widehat{SE}^2 \times \frac{n-1}{n} + \widehat{Bias}^2}$. See section 3.4.1 for details of simulation procedures.

of 95%.

3.4.2 Multiple covariates

One of the main features of the method presented in this Chapter is that it allows for more than one covariate to be easily included in the estimation procedure. To evaluate the performance of the method when more than one covariate is included, we again simulate a population N of 300 objects uniformly distributed in the interval $[0, W]$, with $W = 10$. As before we specify the survey area to be equal to $2LW$, with the transect length $L = 15$ so that the resulting density of objects within the area surveyed, D, is equal to 1.

A total of three covariates were included in the simulations: cluster size, sighting cue, and time of day. Cluster sizes were again generated from a χ^2_5 distribution; values for sighting cue were generated from a discrete Uniform distribution in the interval $(0, 3)$; values for time of day were generated from a Uniform distribution in the interval $(6, 19)$. The detection probabilities associated with each simulated object were obtained based on parameter estimates from a model fitted to offshore spotted dolphin sightings data for the year 1989. The rejection method was then used to determine which objects were detected. Estimates of density D and mean cluster size $\mathbb{E}[s]$ were computed as described for the size-bias case. A total of 240 simulations, with 119 bootstrap iterations carried out for each simulation, were performed, and estimates of bias, standard errors, RMSE and coverage were obtained using the simulation point estimates. Bootstrap iterations in which no convergence was achieved were discarded. Due to the poor performance of the estimator based on a linear form for the scale parameter, simulations were carried out using only the estimator formulated with the scale term as an exponential function of the covariates (Table 3.2).

The point estimate of D showed a much greater bias than those observed for the size-bias case, with very poor coverage. In contrast, estimates of bias and standard error for mean school size $\mathbb{E}[s]$ were comparable to those obtained for the single covariate case, and coverage was close to the nominal value of 95%.

$\hat{D}$	0.9325
RMSE($\hat{D}$)	0.1432
$\widehat{Bias}(\hat{D})$	-0.0675
$\widehat{SE}(\hat{D})$	0.1069
$\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$	-0.6311
Coverage($\hat{D}$)	0.7625
$\hat{\mathbb{E}}[s]$	5.0674
RMSE($\hat{\mathbb{E}}[s]$)	0.3539
$\widehat{Bias}(\hat{\mathbb{E}}[s])$	0.0674
$\widehat{SE}(\hat{\mathbb{E}}[s])$	0.3415
$\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$	0.1975
Coverage($\hat{\mathbb{E}}[s]$)	0.9333

Table 3.2: Estimates of density of clusters ($\hat{D}$) and mean cluster size ($\hat{\mathbb{E}}[s]$) and their corresponding estimates of root mean squared error (RMSE ($\widehat{SE}(\hat{D})$) and RMSE ($\widehat{SE}(\hat{\mathbb{E}}[s])$)), bias ($\widehat{Bias}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])$), standard error ($\widehat{SE}(\hat{D})$ and $\widehat{SE}(\hat{\mathbb{E}}[s])$), the ratio of the estimate of bias to the estimate of standard error ($\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$), and coverage (Coverage($\hat{D}$) and Coverage($\hat{\mathbb{E}}[s]$)) obtained using the method described in section 3.2, with the scale parameter as an exponential function, applied to simulated data. Here the shape ('power') parameter has a value of 2, whereas the scale term is given by the following covariate parameters: 0.0007 for the effect of school size; -1.1578, -1.4524 and -0.9764 for the effects of three levels of sighting cue (with the fourth level corresponding to the intercept); and -0.0062 for the effect of time of day. See section 3.4.2 for details of simulation procedures.

3.5 Example: Applying the method to eastern tropical Pacific dolphin sightings data

In this Section we apply the method to ETP northeastern offshore spotted dolphin sightings data collected by observers placed on board tuna vessels. Buckland and Anganuzzi (1988*b*), Anganuzzi and Buckland (1989) and Anganuzzi *et al.* (1991, 1992) obtained

estimates of $f(0)$ for this stock for the years 1975–1990. We reanalysed these data, investigating the effects of covariates on the detection function and providing estimates of $f(0)$ which take into account any covariate effects. Due to the small sample sizes and few covariate data recorded we do not analyse data from 1975 and 1976. To minimise any potential biases resulting from the data collection procedures employed, especially during the early years, we discarded observations according to the criteria used by Buckland and Anganuzzi (1988*b*), and applied a truncation distance of 5 nm.

A number of variables potentially affect the probability of detection of dolphin schools. Sea state, as indicated by Beaufort, is widely known to affect sightability (e.g. Buckland *et al.* 1993). If weather conditions vary seasonally, then month may also have an effect on detection probabilities. Sighting cues, such as body, splashes, or associated birds, may each represent a different detection function. For example, sightings for which the initial cue was a flock of seabirds flying above the school may be visible over large distances, whereas splashes may not. Sighting method is another potential covariate. Within the last 15 years, the use of helicopters and/or bird radar by tuna vessels to search for schools of dolphins has increased dramatically. The proportion of sightings made by such methods has risen from zero in 1977 to over 33% from 1990 onwards (*cf.* Figure 3.1). It is possible that the detection function for helicopter and/or bird radar sightings may exhibit a broader shoulder than that for sightings made by the crew. Hence we investigate differences in detection probabilities between sightings made by the crew, sightings made by helicopter, and those made by bird radar. A fourth potential covariate is school size, which often affects the detection function. Another covariate to be investigated is time of day, since environmental conditions that vary throughout the day (e.g. the amount of glare) may affect sighting conditions. Finally, we examine whether sightings that led to a set (*i.e.* the vessel approaches the dolphin school and encircles it, retrieving the tuna located underneath the school) exhibited different detection probabilities than those that did not.

We use values for Beaufort ranging from 0 to 3. Although strictly speaking Beaufort is an ordered categorical variable, its relatively large number of parameters resulted in convergence problems, and so we treated it as a continuous variable. Month is also a factor, with 12 levels. However, as with Beaufort, to fit a model with so many parameters was problematic. Hence we reduced the number of levels by contracting the months into four seasons (quarters), according to typical circulation patterns of ETP surface waters (Wyrtki 1966, Fiedler 1992). Thus the first quarter encompasses the months December – February, the second quarter March – May, the third quarter June – August, and the last quarter September – November. Sighting cue was also a factor, with four levels: birds;

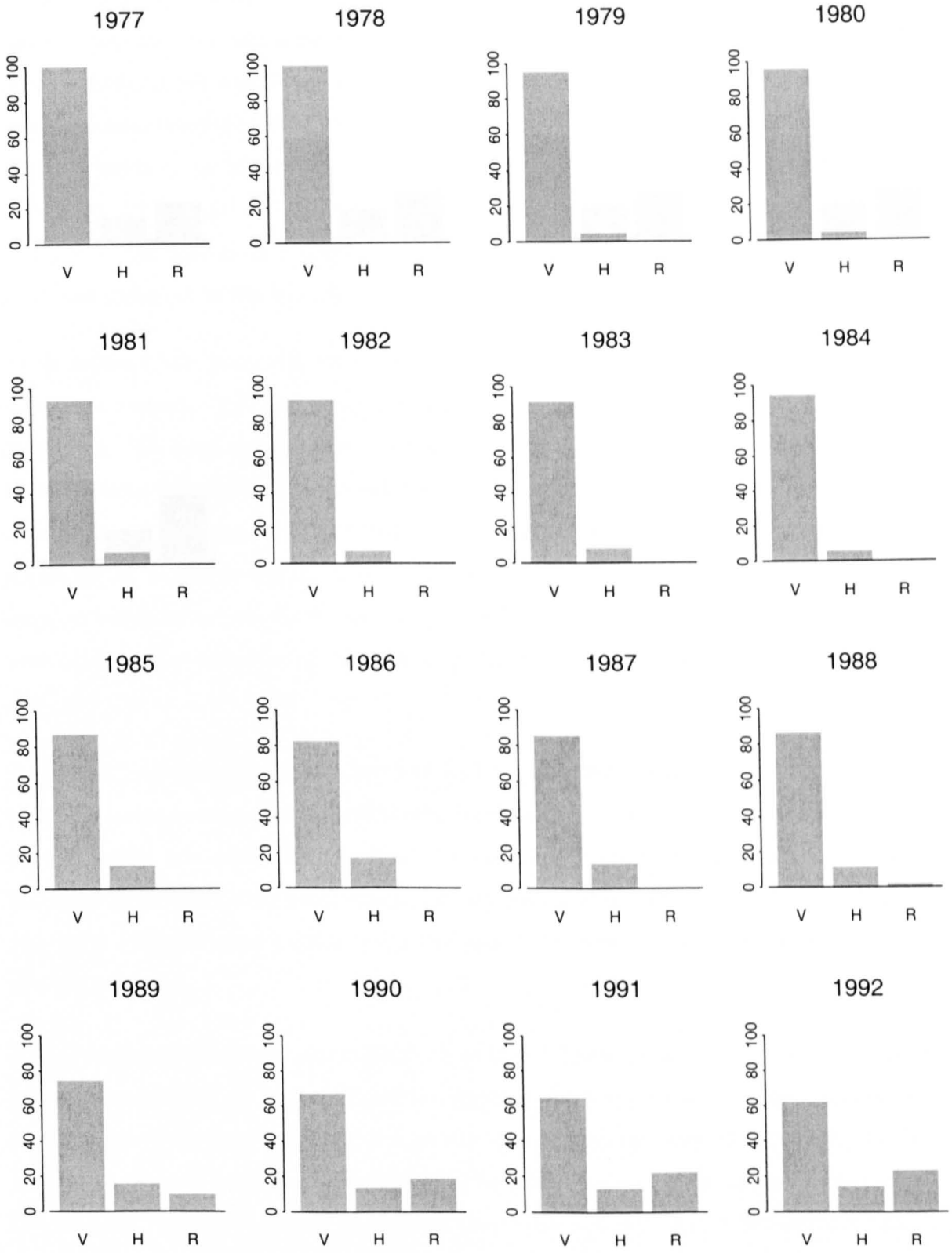


Figure 3.1: Proportion of northeastern offshore spotted dolphin sightings made under different search methods for the years 1977 – 1997. V = sightings made from anywhere on the vessel, by naked eye or using 20x binoculars; H = from helicopter; R = using bird radar.

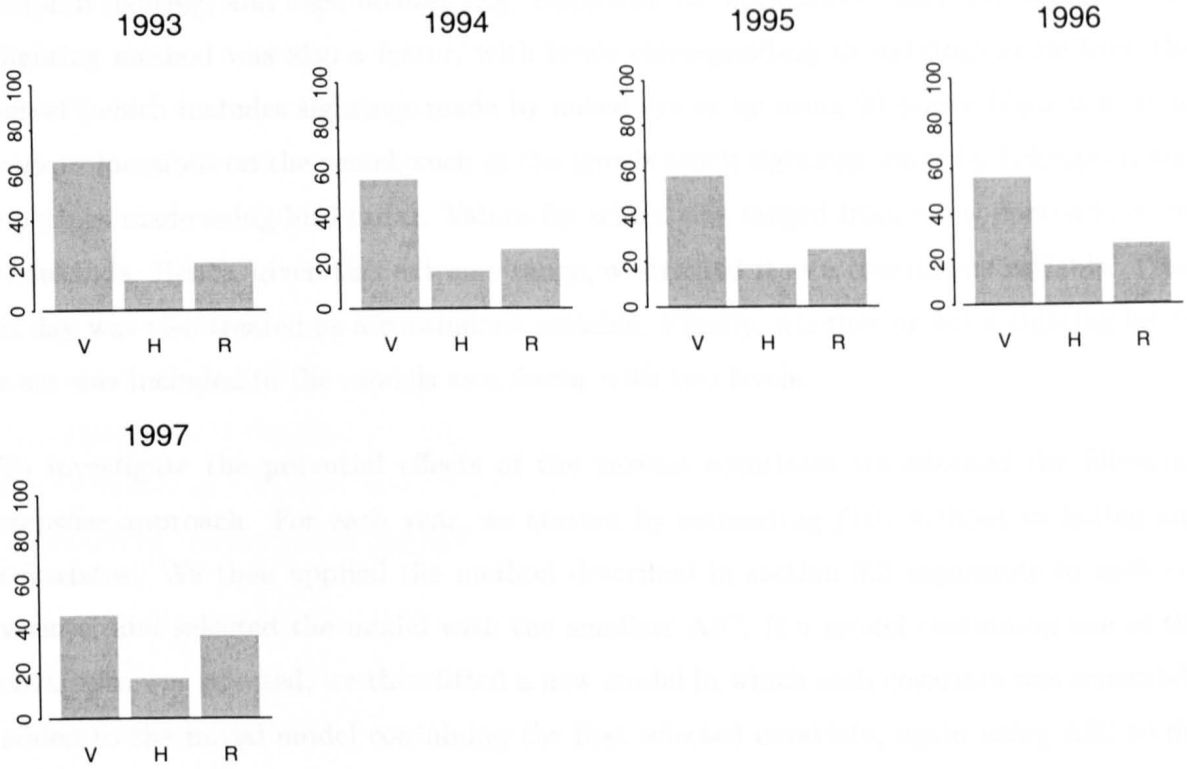


Figure 3.1: Proportion of northeastern offshore spotted dolphin sightings made under different search methods for the years 1977 – 1997. V = sightings made from anywhere on the vessel, by naked eye or using 20x binoculars; H = from helicopter; R = using bird radar.[continued]

splashes; the dolphins themselves or some other cue sighted immediately before the actual dolphin sighting; and logs, breezer (e.g. ripples on the surface) or other non-specified cue. Sighting method was also a factor, with levels corresponding to sightings made from the vessel (which includes sightings made by naked eye or by using 20 power binoculars from various locations on the vessel, such as the crow's nest); sightings made by helicopter; and sightings made using bird radar. Values for school size ranged from a few dozens to a few thousands. Hence, given this extreme range, we treated it as a continuous variable. Time of day was also treated as a continuous variable. Finally, whether or not a sighting led to a set was included in the models as a factor with two levels.

To investigate the potential effects of the various covariates we adopted the following stepwise approach. For each year, we started by estimating $f(0)$ without including any covariates. We then applied the method described in section 3.2 separately to each covariate, and selected the model with the smallest AIC. If a model containing one of the covariates was selected, we then fitted a new model in which each covariate was separately added to the initial model containing the first selected covariate, again using AIC to determine which covariates to retain. We proceeded in this manner until no new covariates were added. An estimate of $f(0)$ for a given year was obtained as the average of the $\hat{f}(0 \mid \mathbf{z})$.

Preliminary analyses indicated that the half-normal model tended to provide better fits than the hazard-rate model, particularly for data from the 1990s. In addition, the half-normal model was more effective than the hazard-rate model in minimising the effect of the spike in the data near zero, which was commonly observed in the data from the 1970s and early 1980s. Hence we only used the half-normal model in all analyses presented in this Chapter.

Initial results yielded poor model fits for a number of years, with unduly large estimates of standard error for $\hat{f}(0)$. Inspection of the data for those years revealed that there were few or no observations for which the cue type was logs, breezer or other non-specified cue. Since those observations amounted to at most 3.5% of all sightings within any year (ranging from 0% to 3.5%, with mean equal to 0.82%), they were not included in subsequent analyses, so that only three levels for cue type were used.

Estimates of $f(0)$, its analytic standard error and coefficient of variation (CV) are presented in Table 3.3. Covariates included in the models for each year are given in Table 3.4. Since one of the main applications of the TVOD is for the assessment of trends in dolphin

relative abundance, the estimates of $f(0)$ presented here can be used to obtain annual estimates of relative abundance. Hence, to minimise the variability in $f(0)$ estimates between years, the covariates most commonly selected in all years were identified and new estimates of $f(0)$ were obtained using only those covariates. Covariates selected in over half of the years were mean school size, cue type, and time of day. However, examination of the models fitted to the data using only those covariates yielded very poor model fits for some years. Of the three covariates, mean school size and cue type were most often the first covariates to be selected (*cf.* Table 3.4), with time of day often coming amongst the last to be included. We then re-fitted the model to the data from all years using only mean school size and cue type as covariates. Although this made little difference to the estimates of $f(0)$ obtained (ranging from -0.26 to 0.01, with mean equal to -0.02)), it considerably improved the model fits (Figure 3.2).

To reduce any bias resulting from the greater amount of effort expended in areas of high dolphin density, estimates of $f(0)$ were initially computed as the average of the estimates for each observation within $f(0)$ strata, where the strata were defined based on the standard procedures of Buckland and Anganuzzi (1988*b*) and Anganuzzi and Buckland (1989). A final estimate of $f(0)$ for each year (referred to as a 'smoothed estimate') was then computed as an area weighted average of the $f(0)$ estimates from all strata (Table 3.3). Results from standard analyses, without incorporating any covariate effects, are also shown for comparison (Table 3.5; Anganuzzi and Buckland 1994, IATTC 1994, 1995, 1996, 1997, 1998, 1999).

For the 1990s, results from analyses with and without covariates are comparable, but for the years prior to about 1992 standard analyses yield considerably larger estimates of $f(0)$ (Table 3.5), suggesting a strong trend of increase in estimates of the effective strip (half-)width (the inverse of $f(0)$) over time. Data from the early years (1970s to mid-1980s) contain a large number of schools detected on the transect line (Figure 3.3), a result of the vessels having turned towards the schools before the observer recorded the original sighting angles (Buckland and Anganuzzi 1988*b*). The hazard-rate model used in the standard analyses appears to overestimate $f(0)$ as it tries to fit the spike in the data at distance zero. As the extent of rounding of perpendicular distances to zero decreases over time, so do the $f(0)$ estimates, and hence the observed trend of increase in effective strip (half-)width. Replacing the hazard-rate model by the half-normal model minimises this effect (Lennert-Cody *et al.*, In prep.). The inclusion of covariates, in addition to the use of the half-normal model, appears to further reduce the trend in $f(0)$ estimates over time, but this effect is small.

Year	$\hat{f}_c(0)$	$\widehat{SE}_c\{\hat{f}(0)\}$	$\%CV_c\{\hat{f}(0)\}$	$\hat{f}_s(0)$	$\widehat{SE}_s\{\hat{f}(0)\}$	$\%CV_s\{\hat{f}(0)\}$
1977	0.2639	0.0298	11.31	0.2877	0.0048	1.69
1978	0.3177	0.0136	4.29	0.3204	0.0061	1.90
1979	0.2779	0.0222	7.98	0.2784	0.0065	2.32
1980	0.3095	0.0131	4.22	0.3110	0.0044	1.42
1981	0.2929	0.0125	4.28	0.2926	0.0049	1.68
1982	0.3075	0.0124	4.03	0.3088	0.0044	1.42
1983	0.3166	0.0150	4.72	0.3108	0.0072	2.33
1984	0.2971	0.0481	16.20	0.2936	0.0098	3.35
1985	0.3249	0.0077	2.36	0.3230	0.0034	1.04
1986	0.3074	0.0090	2.90	0.3063	0.0037	1.20
1987	0.2878	0.0070	2.40	0.2917	0.0021	0.72
1988	0.2858	0.0206	7.19	0.2809	0.0025	0.91
1989	0.2522	0.0069	2.75	0.2626	0.0022	0.84
1990	0.2616	0.0078	2.98	0.2693	0.0030	1.11
1991	0.2683	0.0207	7.73	0.2662	0.0038	1.43
1992	0.2716	0.0083	3.08	0.2688	0.0025	0.92
1993	0.2617	0.0175	6.68	0.2639	0.0037	1.40
1994	0.2567	0.0123	4.80	0.2588	0.0036	1.40
1995	0.2507	0.0069	2.74	0.2518	0.0031	1.24
1996	0.2392	0.0055	2.32	0.2383	0.0047	1.99
1997	0.2668	0.0079	2.96	0.2650	0.0032	1.22

Table 3.3: Estimates of $f(0)$ ($\hat{f}(0)$), its analytic standard error ($\widehat{SE}\{\hat{f}(0)\}$) and percent coefficient of variation ($\%CV\{\hat{f}(0)\}$) for northeastern offshore spotted dolphins for the years 1977 – 1997. Column headings with the subscript c denote estimates obtained from analyses with covariates; those with the subscript s indicate estimates based on an area-weighted average of $f(0)$ estimates obtained for each of the strata determined via the post-stratification method.

Year	Month	Beaufort	Time of day	School size	Set	Sighting cue	Method
1977		2	3	1			
1978					1		
1979			1	2	3		
1980						1	
1981	5	4	3	2		1	
1982				1		2	
1983					1		
1984	3	5	1		4	2	
1985		2	1				
1986	2	3		1	4		
1987		2	1		3		
1988		2	5		1	3	4
1989			1	2			
1990	4		1	3		2	
1991			4		3	2	1
1992	4			1		2	3
1993		2	3	1			
1994	3	4	2	1			
1995	5	2	3	1		4	
1996				1	3	2	
1997	4			3		2	1
Total	8	10	13	13	9	11	4

Table 3.4: Selected covariates for models applied to northeastern offshore spotted dolphin sightings data for the years 1977 – 1997. The number indicates the order in which covariates were selected. Also shown are the total number of times each covariate was selected in all years.

Year	$\hat{f}_b(0)$	$\widehat{SE}_b\{\hat{f}(0)\}$	$\%CV_b\{\hat{f}(0)\}$	$\hat{f}_o(0)$	$\widehat{SE}_o\{\hat{f}(0)\}$	$\%CV_o\{\hat{f}(0)\}$
1977	0.2958	0.0188	6.36	0.3636	0.0463	12.73
1978	0.3289	0.0203	6.18	0.4484	0.0603	13.45
1979	0.2903	0.0268	9.22	0.4202	0.0530	12.61
1980	0.3143	0.0187	5.95	0.4016	0.0500	12.45
1981	0.2975	0.0150	5.03	0.3650	0.0360	9.85
1982	0.3148	0.0148	4.70	0.3802	0.0492	12.93
1983	0.3207	0.0197	6.15	0.3049	0.0325	10.67
1984	0.3036	0.0228	7.50	0.3367	0.0363	10.77
1985	0.3199	0.0129	4.03	0.4065	0.0297	7.32
1986	0.3106	0.0124	4.00	0.3937	0.0310	7.87
1987	0.2929	0.0066	2.26	0.3436	0.0154	4.47
1988	0.2846	0.0081	2.84	0.3145	0.0237	7.55
1989	0.2664	0.0087	3.27	0.3040	0.0222	7.29
1990	0.2740	0.0093	3.40	0.2976	0.0159	5.36
1991	0.2729	0.0104	3.81	0.2899	0.0143	4.93
1992	0.2753	0.0061	2.22	0.2825	0.0128	4.52
1993	0.2700	0.0081	2.99	0.2519	0.0076	3.02
1994	0.2700	0.0176	6.52	0.2545	0.0084	3.31
1995	0.2579	0.0064	2.48	0.2532	0.0077	3.04
1996	0.2488	0.0139	5.57	0.2469	0.0110	4.44
1997	0.2717	0.0069	2.55	0.2597	0.0074	2.86

Table 3.5: Estimates of $f(0)$ ($\hat{f}(0)$), its analytic standard error ($\widehat{SE}\{\hat{f}(0)\}$) and percent coefficient of variation ($\%CV\{\hat{f}(0)\}$) for northeastern offshore spotted dolphins for the years 1977 – 1997. Column headings with the subscript b denote bootstrap estimates obtained from analyses with covariates; those with the subscript o indicate original estimates without covariates.

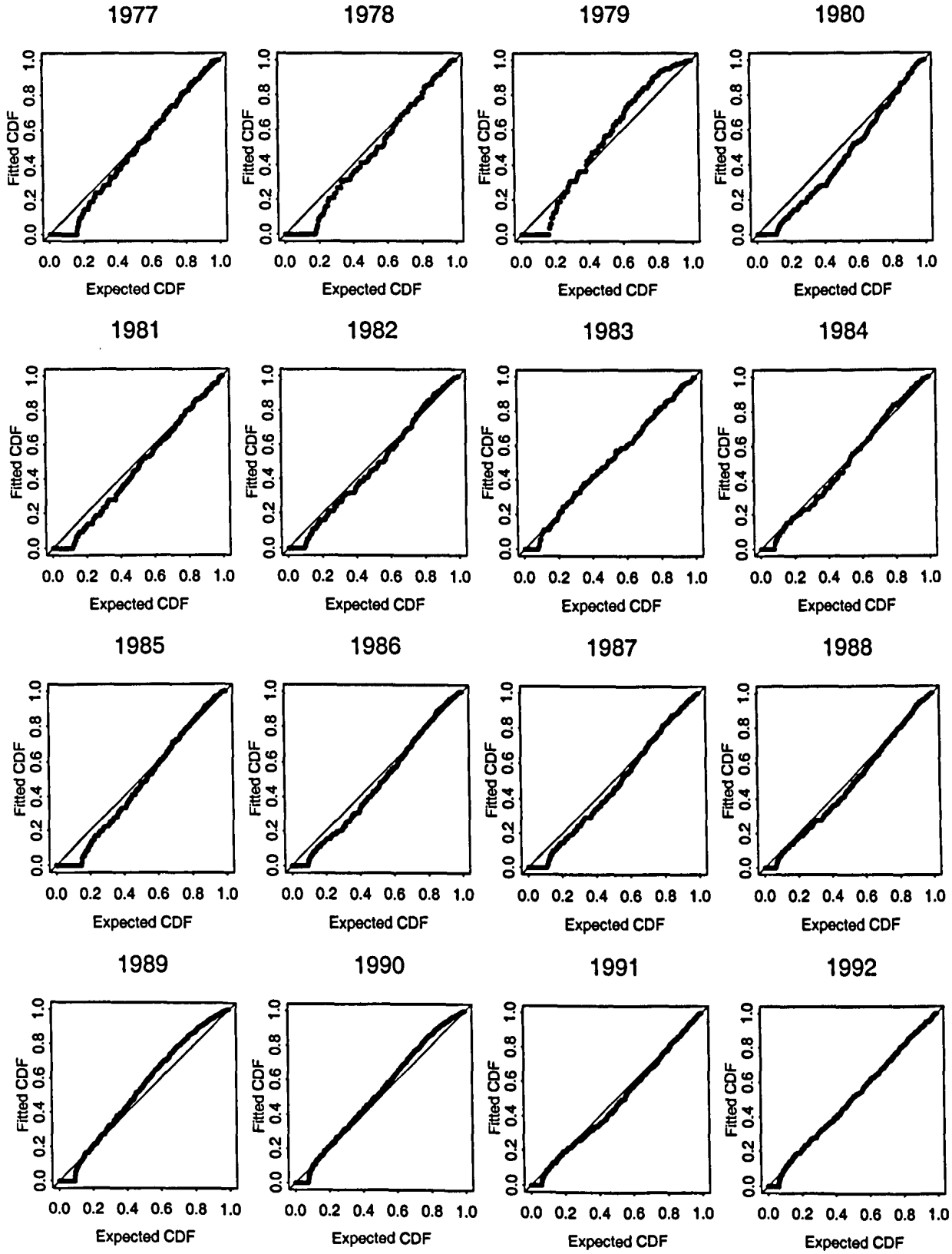


Figure 3.2: qq-plots of the cumulative density function (CDF) of the estimates of $f(0)$ for northeastern offshore spotted dolphin sightings data for the years 1977 - 1997.

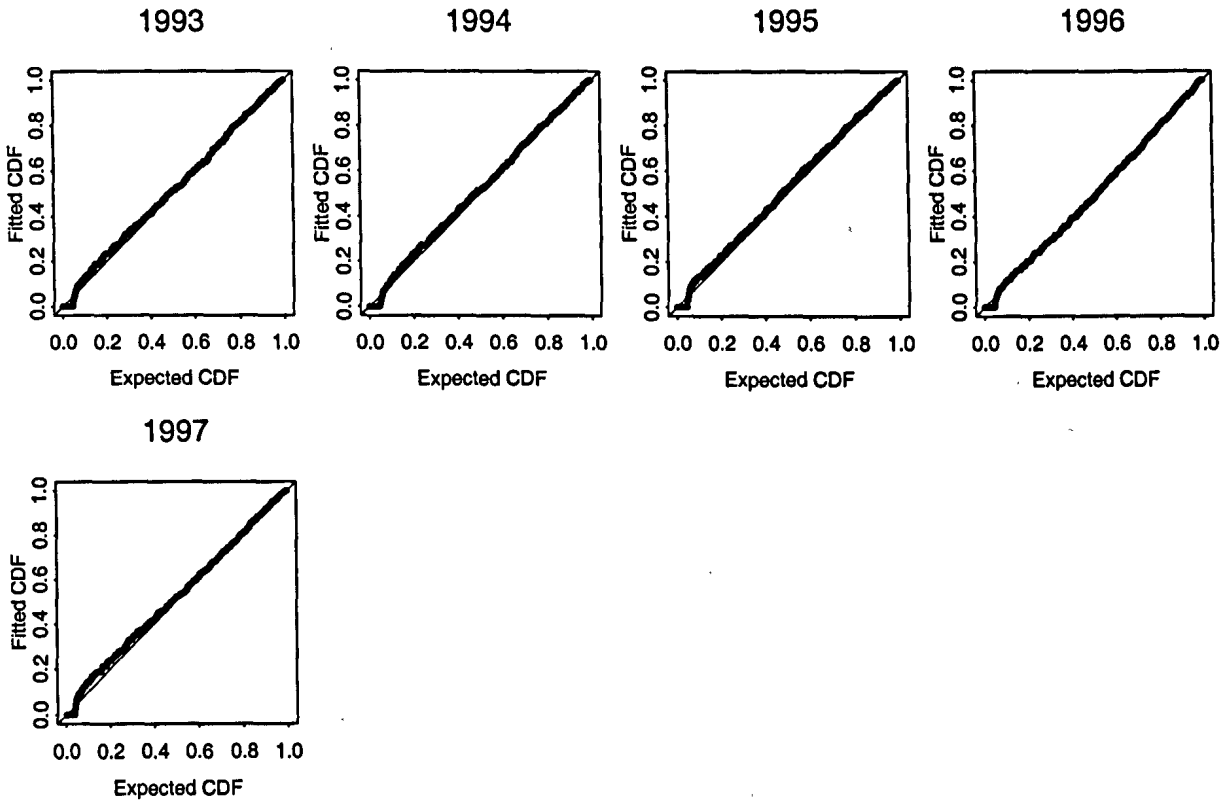


Figure 3.2: qq-plots of the cumulative density function (CDF) of the estimates of $f(0)$ for northeastern offshore spotted dolphin sightings data for the years 1977 – 1997. [continued]

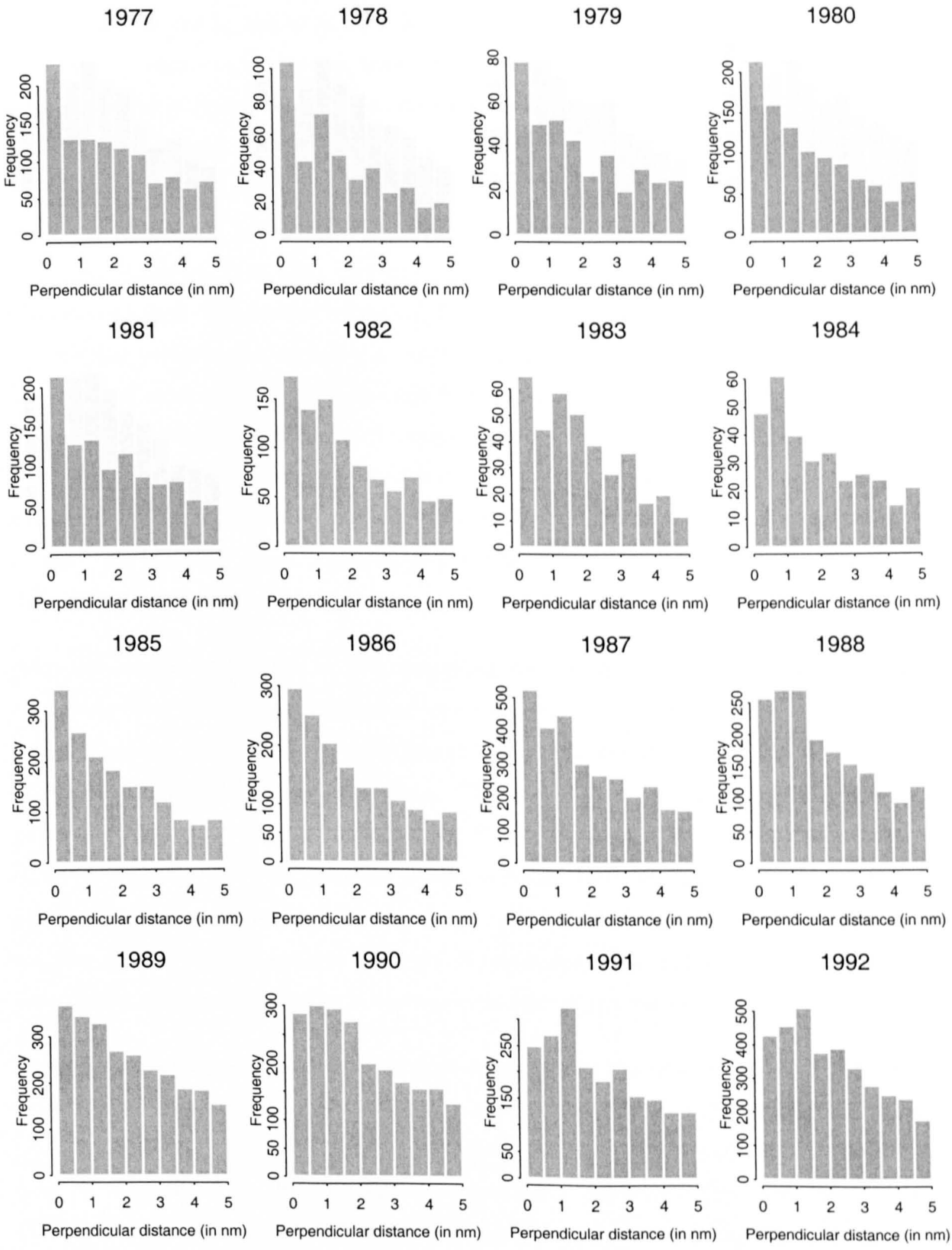


Figure 3.3: Histograms of perpendicular distances of offshore spotted dolphin sightings for the years 1977 – 1997.

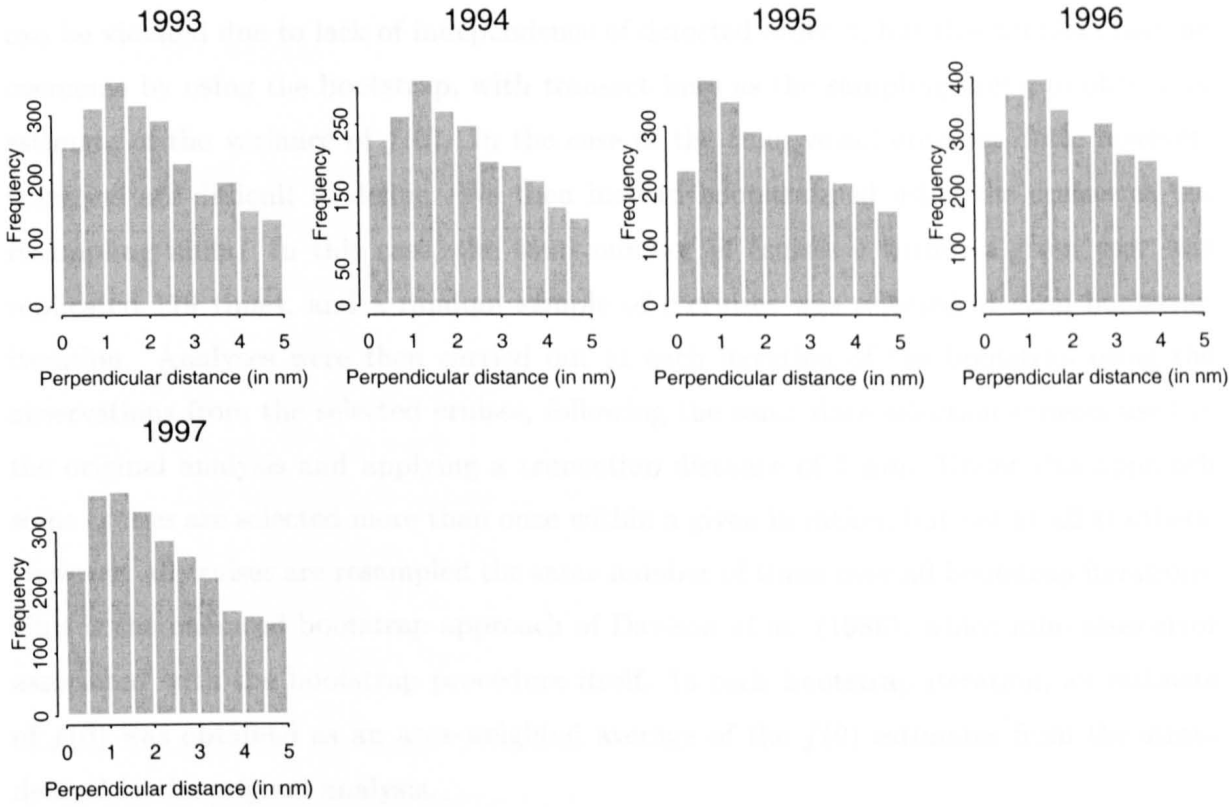


Figure 3.3: Histograms of perpendicular distances of offshore spotted dolphin sightings for the years 1977 – 1997. [continued]

Although analytic estimates of the variance may be obtained, they rely on the assumption of asymptotic normality of the maximum likelihood estimates of $f(0 \mid z)$. This assumption can be violated due to lack of independence of detected objects, but this problem may be overcome by using the bootstrap, with transect lines as the sampling units, to obtain an estimate of the variance of $\hat{f}(0)$. In the case of the tuna vessel observer data, however, transects are difficult to define. We then instead bootstrapped using the cruises as the resampling units. In this case, the total number of cruises c within a given year was replicated 119 times, and a random sample of c cruises was selected at each bootstrap iteration. Analyses were then carried out at each iteration of the bootstrap using the observations from the selected cruises, following the same data selection criteria used in the original analysis and applying a truncation distance of 5 nm. Under this approach some cruises are selected more than once within a given iteration, but not at all at others. However, all cruises are resampled the same number of times over all bootstrap iterations. This is the balanced bootstrap approach of Davison *et al.* (1986), which minimises error associated with the bootstrap procedure itself. In each bootstrap iteration, an estimate of $f(0)$ was obtained as an area-weighted average of the $f(0)$ estimates from the strata defined in the original analysis.

Although cruises were taken as the resampling units when bootstrapping, which implies that detections between cruises are independent, this may not necessarily be the case. Tuna vessels in the ETP form 'code groups', with the aim of maximising their chances of finding dolphin schools carrying tuna. This means that once a vessel belonging to a particular code group finds dolphin schools, it then informs other vessels belonging to the same code group about the location of such schools. However, the extent and nature of the information exchanged within and between code groups are extremely complex, and unlikely to follow any consistent patterns. Nonetheless, the existence of such code groups may introduce some bias in the bootstrap estimates of the variance of $\hat{f}(0)$.

Table 3.5 shows bootstrap estimates based on cruises as the resampling units. Bootstrap estimates of standard error were generally greater than the analytic estimates (*cf.* Table 3.3), suggesting that the latter were negatively biased. It is possible that such bias may be a result of the post-stratification procedure applied to the data. If estimates of $f(0)$ within each stratum are generally similar, then the resulting estimates of their standard error will be small. Since bootstrap estimates of standard error are based on the final estimates of $f(0)$ from all iterations, they are not affected by any bias caused by the post-stratification procedure. Bootstrap coefficients of variation for estimates up until the 1990s were smaller than those from the original estimates, becoming somewhat

comparable in the late 1990s (Table 3.5).

3.6 Discussion

Simulation results for size-biased data were consistent for all combinations of parameter values, suggesting that the method is robust to varying extents of size-bias effects. In contrast, simulation results for the model containing multiple covariates yielded poor coverage, and also greater bias, for estimates of density. In a number of simulations, the bootstrap estimates of density fell either all below, or all above, 1, resulting in the poor coverage observed. It is possible that the relatively narrow range of the bootstrap estimates of density within each simulation may have been a result of convergence problems. As the number of parameters in the model is increased, the more likely it is that convergence problems will occur. This could explain the difference in simulation results for the single and multiple covariate cases.

For the simulations, bootstrap iterations in which the models failed to converge were discarded, with additional iterations carried out until a total of 119 bootstrap estimates were obtained. For the size-bias case, this resulted in a mean of approximately 3.5% of the bootstrap iterations within a given simulation being discarded. For the multiple covariate case, the percentage of iterations which failed to converge was much greater, with the mean at around 30%. However, if only one of the two contending models (half-normal or hazard-rate) failed to converge, the other model - which converged - was selected, although it may not have been an appropriate choice. Hence we are likely introducing some bias in the simulation results for multiple covariates, given its greater incidence of lack of convergence. To investigate this, a total of 240 simulations, with 119 bootstrap iterations, were carried out, allowing for different covariates to be selected at each iteration. Simulations were separately run using either the half-normal or the hazard-rate model. A substantial improvement in the coverage for the estimates of density was obtained, as well as a reduction in bias (Table 3.6).

Simulation results for the size-bias scenario compared favourably with those obtained by Chen (1996) using a kernel method. Unlike Chen's estimator, coverage of the density estimator based on the conditional likelihood approach was consistently close to the nominal value of 95% under all combinations of parameter values. In addition, the bias was smaller than that obtained using the kernel method. Similar results were obtained for the mean school size estimates.

	Half-normal model	Hazard-rate model
$\hat{D}$	1.0220	0.9923
RMSE($\hat{D}$)	0.1044	0.1179
$\widehat{Bias}(\hat{D})$	0.0220	-0.0077
$\widehat{SE}(\hat{D})$	0.1023	0.1184
$\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$	0.2151	-0.0650
Coverage($\hat{D}$)	0.9042	0.9542
$\hat{\mathbb{E}}[s]$	5.0870	5.1060
RMSE($\hat{\mathbb{E}}[s]$)	0.3242	0.3306
$\widehat{Bias}(\hat{\mathbb{E}}[s])$	0.0870	0.1060
$\widehat{SE}(\hat{\mathbb{E}}[s])$	0.3130	0.3479
$\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$	0.2778	0.3047
Coverage($\hat{\mathbb{E}}[s]$)	0.9625	0.9708

Table 3.6: Estimates of density of clusters ($\hat{D}$) and mean cluster size ($\hat{\mathbb{E}}[s]$) and their corresponding estimates of root mean squared error (RMSE ($\widehat{SE}(\hat{D})$) and RMSE ($\widehat{SE}(\hat{\mathbb{E}}[s])$)), bias ($\widehat{Bias}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])$), standard error ($\widehat{SE}(\hat{D})$ and $\widehat{SE}(\hat{\mathbb{E}}[s])$), the ratio of the estimate of bias to the estimate of standard error ($\widehat{Bias}(\hat{D})/\widehat{SE}(\hat{D})$ and $\widehat{Bias}(\hat{\mathbb{E}}[s])/\widehat{SE}(\hat{\mathbb{E}}[s])$), and coverage (Coverage($\hat{D}$) and Coverage($\hat{\mathbb{E}}[s]$)) obtained using the method described in section 3.2, with the scale parameter as an exponential function, applied to simulated data. Here the shape ('power') parameter has a value of 2, whereas the scale term is given by the following covariate parameters: 0.0007 for the effect of school size; -1.1578, -1.4524 and -0.9764 for the effects of three levels of sighting cue (with the fourth level corresponding to the intercept); and -0.0062 for the effect of time of day. See section 3.6 for details of simulation procedures.

In standard analyses of line transect data, model selection is aided by the use of goodness of fit tests (Buckland *et al.* 1993). Although for the method presented in this Chapter relative measures of model fit (e.g. AIC) can be employed to select among competing models, standard tests for evaluating model fit cannot be applied. Likelihood ratio tests,

for example, require knowledge of the distribution of the multi-dimensional detection function. Although the distribution of the perpendicular distances is implicitly assumed to be uniform in the interval (0,W), the distribution of the covariates in the population is usually not known, and so would have to be assumed. If the assumed distributions are incorrect, test results could be misleading. Alternatively, Chi-square tests would not rely on any distributional assumptions, but depending on the number of covariates used in the models, the required grouping of observations would render the test impractical. However, the appropriateness of a given model may be informally evaluated through the use of diagnostic plots. An example would be theoretical qq-plots (*cf.* Chambers *et al.* 1983, Cleveland 1993), in which quantiles of the data are plotted against quantiles of a theoretical distribution hypothesized to correspond to that of the data. If we take the data $\mathbf{Y}$ ($\mathbf{Y} = Y_1, \ldots, Y_n$) to be the estimates of $f(x_i \mid \mathbf{z}_i)$ obtained for each observation, denote by $\hat{F}(Y_i)$ the estimated cumulative density function (CDF) for each Y_i, which is given by $\int_0^{x_i} \hat{f}(x \mid \mathbf{z}_i)\, dx$. From the Probability-Integral Transform Theorem the CDFs of independent and identically distributed random variables $F_0(\mathbf{Y})$ ($F_0(\mathbf{Y}) = F_0(Y_1), \ldots, F_0(Y_n)$) are uniformly distributed on the interval $(0, 1)$. Hence, if the estimated model for the detection function fits the data well, $\hat{F}(\mathbf{Y}) = F_0(\mathbf{Y})$ and so $(\hat{F}(Y_1), \ldots, \hat{F}(Y_n)) \sim (U_1(0,1), \ldots, U_n(0,1))$. By plotting the values of our fitted CDFs against a $U(0,1)$ distribution we can assess how well our estimated model fits the data. This approach is equivalent to the standard plotting of observed versus fitted values, in which plotted observations are expected to follow a straight line if the fitted values agree well with the data.

Figure 3.2 shows theoretical qq-plots for the estimates of $f(x \mid \mathbf{z})$ for the tuna vessel observer data for the years 1977 – 1997. Note that, in all plots, the straight line at a value of zero for the fitted CDF corresponds to detections made on the line (i.e. at a perpendicular distance of zero), so that the fitted CDF is also equal to zero. This feature of the plots highlights the unduly large number of detections on the line that were made during the 1970s and early 1980s. In general the estimated CDFs for the covariate models approximate the theoretical CDFs relatively well.

Unlike methods based on double-platform data (e.g. Borchers *et al.* 1998*a*), standard line transect data do not allow estimation of the distribution of the covariates in the population. Hence our plots of the detection function for any given covariate (e.g. Figure 3.4) are conditional on values of the other covariates included in the model. When the other covariates are continuous, we use values of those covariates corresponding to their 25th, 50th and 75th quantiles. In the case of factors, we use all combinations of the levels of any factors included in the model. This way the interpretation of the effects of the various

covariates is facilitated.

A desired property of any model for the detection function is that it is 'pooling robust', that is, it yields reliable estimates even when data are pooled over many covariates which may affect the detection function (Buckland *et al.* 1993). Most standard line transect estimators are pooling robust. However, when strong covariate effects are present, the method presented in this Chapter provides an alternative to stratification. It is particularly useful for opportunistic survey data, which commonly have poor spatial coverage and small sample sizes. However, data collected from opportunistic survey platforms may still result in biased estimates of $f(0)$, and hence of density, even when covariate effects are taken into account. The tuna vessel observer data provide an example of an opportunistic survey platform, in which search effort is not random relative to the distribution of the species of interest. In the ETP, effort is concentrated in areas of high dolphin densities, so that the vessels can maximise their chance of catching the large schools of tuna commonly associated with some species of dolphins. In areas of lower dolphin densities, where the intensity of effort is also lower, other cues such as logs or other flotsam are also used for locating the tuna. Because in such areas dolphin schools are not the only target, search for dolphins is not as actively pursued as in areas of high dolphin densities. Hence the effective strip (half-)width in areas of lower dolphin densities is likely to be smaller than that for the high density areas. As the majority of sightings are made in the high density areas, standard analysis of these data, regardless of whether any covariate effects are taken into account, may result in an underestimate of $f(0)$ (i.e. overestimate of the effective strip (half-)width).

In addition to bias in $f(0)$ estimates, standard abundance estimation procedures using opportunistic survey data will also result in biases in encounter rate and mean school size estimates, as the population sampled is not representative of the entire survey region. In the case of the tuna vessel observer data, where vessels concentrate their effort in areas of high dolphin densities and also target larger schools, which potentially have larger schools of tuna associated with them, both encounter rate and mean school size will be overestimated.

There are two ways in which the resulting bias in estimates of dolphin abundance based on opportunistic survey data may be minimised, using the estimation procedure described in this Chapter. One way is by using post-stratification to separately obtain estimates of encounter rate, mean school size and $f(0)$ (Buckland and Anganuzzi 1988*b*, Anganuzzi and Buckland 1989). Under this approach, the survey region is divided into a grid of cells.

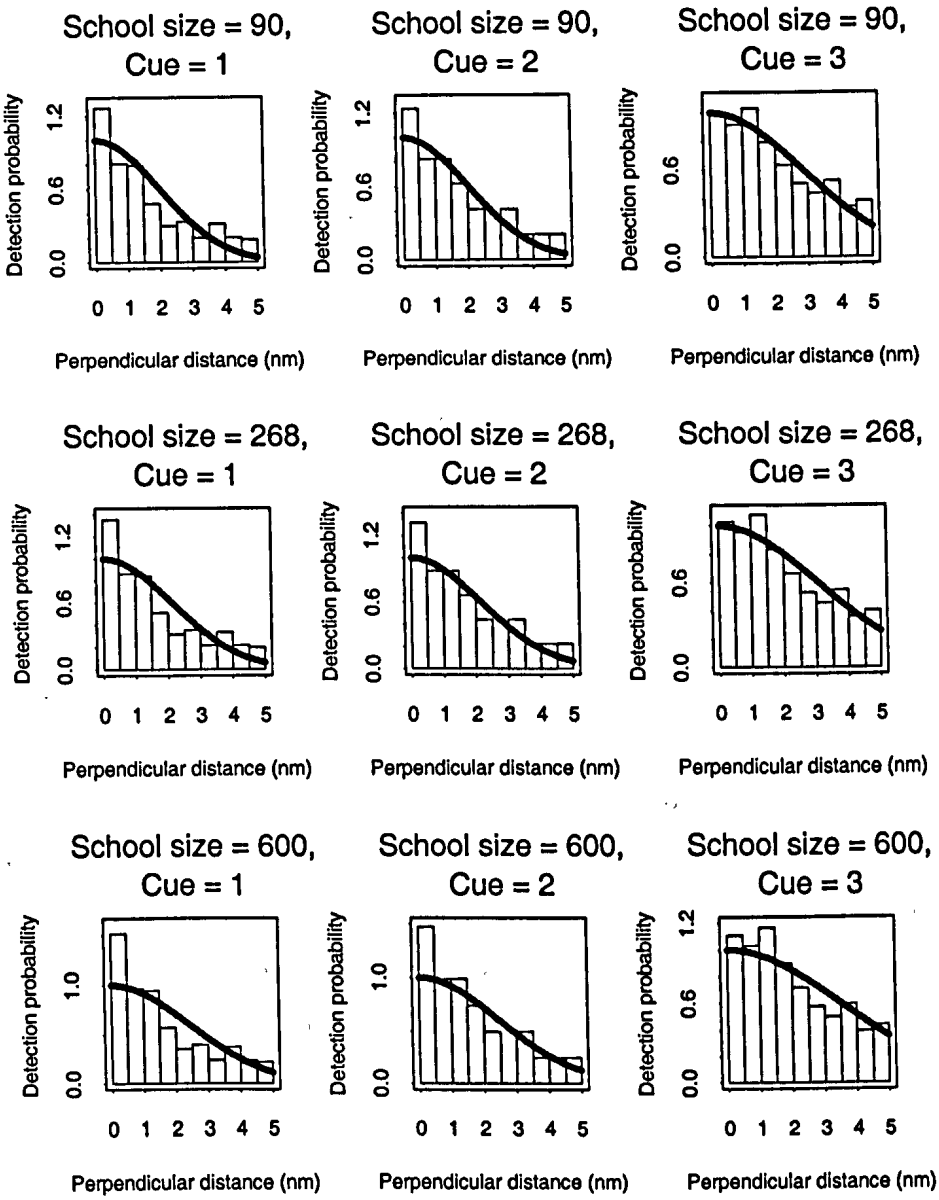


Figure 3.4: Histogram of perpendicular distances and fitted detection function for levels of the covariate cue type based on tuna vessel observer data from 1982. Cue 1 = dolphins; Cue 2 = splashes; Cue 3 = birds.

If sample size is large enough, cells may be defined both in space and time (e.g. month or year); otherwise they will represent spatial location only, that is, data are pooled over months, years, or some other temporal scale. Estimates of encounter rate, mean school size and $f(0)$ are then obtained for each cell for which there are observations. A smoothing procedure can then be applied to generate values for encounter rate, mean school size and $f(0)$ over cells which were not surveyed. Strata may then be defined based on similarities between values of the quantity of interest (e.g. encounter rate) along adjacent cells. In the past, however, estimates of $f(0)$ could not be obtained for each cell due to small sample sizes within cells. Instead, strata were defined based on some other variable which correlated well with $f(0)$ or its inverse (effective strip (half-)width). However, using the approach described in this Chapter, we can now obtain estimates of $f(0 \mid \mathbf{z}_{ilt})$ for each observation i ($i = 1, \ldots, n$) within cell l ($l = 1, \ldots, m$) at time t ($t = 1, \ldots, T$) based on the estimated parameters from the model fitted to the pooled observations, and an estimate of $f_{lt}(0)$ for that cell can be obtained as the average of the $\hat{f}(0 \mid \mathbf{z}_{ilt})$. The smoothing procedure can then be applied, and strata defined based on the smoothed estimates of $f_{lt}(0)$. Final estimates of encounter rate, mean school size and $f(0)$ can then be obtained as an average of the estimates from each stratum, weighted by their area. These final estimates may then be combined to yield an overall estimate of abundance. An estimate of the variance of the abundance estimate may then be obtained by bootstrapping, using the cruises as resampling units, and repeating the analysis described above at each iteration. However, a limitation of this approach is that mean school size cannot be included as a covariate in the model for $f(0)$.

A second approach is to employ the spatio-temporal modelling framework described in Chapter 2, as follows. We start by using the estimated model parameters to compute estimates of $f(0 \mid \mathbf{z}_{ilt})$, as previously described. The Horvitz-Thompson-like estimator of expression (3.39) can then be applied to yield an estimate of abundance $\hat{N}_{lt}$ for cell lt. Not all $m \times T$ cells will contain observations, although it is assumed that a set of explanatory variables is available for all cells. We then compute $\hat{N}_{lt}$ for cells which have been surveyed. To obtain an estimate of abundance for the entire survey region, we can then model the $\hat{N}_{lt}$ as a function of the spatially and temporally explicit explanatory variables $\mathbf{u}$, and integrate under the predicted surface. Although it is possible to model the $\hat{N}_{lt}$ directly, in practice the area of each cell varies with latitude. Hence it is convenient to model the estimated cell densities ($\hat{D}_{lt}$) instead, with the cell areas (A_l) as an offset in the model. Using the generalised additive model framework of Chapter 2, we then have:

$$\hat{D}_{lt} = \exp\left\{\log_e(A_l) + \beta_0 + \sum_{k=1}^{q} f_k(u_{klt})\right\} + \epsilon_{lt}, \tag{3.45}$$

where u_{klt} $(k = 1, \ldots, q)$ corresponds to the kth explanatory variable available for each cell l at time t, $f_k(\cdot)$ is a smooth function of the kth covariate u_k, and ϵ_{lt} is the error term. A Gamma error distribution may be used (e.g. Augustin 1999). The variance may be estimated by bootstrapping on cruises and repeating the procedure described above at each iteration. The advantage of this approach is that mean school size can then be included as a covariate in the models for $f(0)$.

3.7 Appendices

3.7.1 Derivation of $f(x, \mathbf{z})$

The derivation of the multivariate pdf $f(x, \mathbf{z})$ follows that from Seber (1982) for the univariate case. Define:

$$\Pr[\text{object at } x] = \pi(x)$$

$$\Pr[\text{object has covariate values } \mathbf{z}] = \pi(\mathbf{z})$$

$$\Pr[\text{object detected} \mid \text{object at } x \text{ and has covariate values } \mathbf{z}] = g(x, \mathbf{z}).$$

Assuming that the x's and $\mathbf{z}$'s are independent, then the joint density for the x and $\mathbf{z}$ in the population, $\pi(x, \mathbf{z})$, corresponds to the product of the individual densities $\pi(x)$ and $\pi(\mathbf{z})$. Then:

$$\Pr[\text{object detected at } x \text{ and has covariate values } \mathbf{z}] = g(x, \mathbf{z})\ \pi(x)\ \pi(\mathbf{z})$$

and

$$\Pr[\text{object detected}] = \int_{\mathbf{Z}} \int_{X} g(x, \mathbf{z})\ \pi(x)\ \pi(\mathbf{z})\ dx\ d\mathbf{z},$$

so that:

$$\Pr[\text{object at } x \text{ and has covariate values } \mathbf{z} \mid \text{object detected}]$$

$$= \frac{\Pr[\text{object detected at } x \text{ and has covariate values } \mathbf{z}]}{\Pr[\text{object detected}]}$$

$$= \frac{g(x,\mathbf{z})\ \pi(x)\ \pi(\mathbf{z})}{\int_{\mathbf{Z}} \int_X g(x,\mathbf{z})\ \pi(x)\ \pi(\mathbf{z})\ dx\ d\mathbf{z}}$$

$$= f(x,\mathbf{z}).$$

3.7.2 Generalising the estimator to point transects and to grouped (line transect or point transect) data

In the context of point transects we have detection distances r, with $\pi(r) = r/W$, $0 \leq r \leq W$. Substituting r and $\pi(r)$ for x and $\pi(x)$ in expression (3.2) gives:

$$f(r,\mathbf{z}) = \frac{r\, g(r,\mathbf{z})\, \pi(\mathbf{z})}{\int_{\mathbf{Z}} \int_X r\, g(x,\mathbf{z})\, \pi(\mathbf{z})\, dx\, d\mathbf{z}}.$$

Following the derivation from section 3.1, the conditional likelihood is then given by:

$$\mathcal{L}(\theta;\mathbf{r},\mathbf{z}) = \prod_{i=1}^{n} \frac{r_i\, g(r_i,\mathbf{z}_i)}{\int_R r\, g(r,\mathbf{z}_i)\ dr}.$$

Applying the above result to expression (3.9), and replacing $k'(r,\mathbf{z}) = r \cdot k(r,\mathbf{z})$ for $k(r,\mathbf{z})$, parameter estimates can then be obtained as described for the line transect case.

In the case of grouped data, let $x_{j,i}$ ($j = 1,\ldots,m$, $i = 1,\ldots,n$) denote the ith observed perpendicular distance within the jth distance interval in which the data were grouped, and let j be chosen such that $x_{j-1,i} < x_i \leq x_{j,i}$. Similarly let $r_{j,i}$ be the equivalent for point transects. For line transect data the conditional likelihood is then given by:

$$\mathcal{L}(\theta; \mathbf{x}, \mathbf{z}) = \prod_{i=1}^{n} f(x_i \mid \mathbf{z}_i)$$

$$= \prod_{i=1}^{n} \frac{\int_{x_{j-1,i}}^{x_{j,i}} g(x_i, \mathbf{z}_i)\, dx}{\int_X g(x, \mathbf{z}_i)\, dx}$$

and, for point transects:

$$\mathcal{L}(\theta; \mathbf{r}, \mathbf{z}) = \prod_{i=1}^{n} f(r_i \mid \mathbf{z}_i)$$

$$= \prod_{i=1}^{n} \frac{\int_{r_{j-1,i}}^{r_{j,i}} r\, g(r, \mathbf{z}_i)\, dr}{\int_R r\, g(r, \mathbf{z}_i)\, dr}.$$

Chapter 4

Estimating trends in abundance - an alternative approach

4.1 Introduction

A variety of definitions for what constitutes a "stock" have been put forward since the idea of separately managing populations of a given species was first presented at the beginning of the 20*th* century (see Gauldie (1991) and Dizon *et al.* (1992) for a review). Currently favoured conceptual definitions generally focus on the value of the genetic uniqueness of locally adapted populations. Since local adaptation is an evolutionary process, implicit in these definitions is a certain degree of reproductive isolation between populations. As it is not feasible to directly measure rates of gene flow between populations, a number of proxies are commonly used to infer reproductive isolation, and hence to define stocks.

When abundance estimation procedures are applied to stocks of a given species, the reliability of the resulting estimates depends on the assumptions that (i) stocks have been correctly identified; (ii) stock boundary locations have been correctly determined; and (iii) movement of individuals between stocks is negligible or non-existent. Management decisions based on incorrect stock classifications can potentially lead to the extinction of locally adapted populations, or to a contraction of their (historical) range.

To minimize these risks, assessment methods that are robust to assumptions about the existence and/or location of stock boundaries are desirable. Such an approach has been implemented by the International Whaling Commission (IWC) to the management of large whales, where the traditional management by stock has been replaced by the concept of Small Management Areas (SMAs; IWC 1993*b*). SMAs consist, by default, of 10° × 10°

latitude–longitude blocks where one or more stocks may be present (IWC 1992, 1993*a*); the size of an SMA may vary if there is enough evidence to justify it. Because there are no *a priori* assumptions about stock structure within an SMA, stock status assessments are based on the observed proportion of a given species within the SMA. In addition, estimates of abundance are separately obtained for each SMA, thus avoiding the risks of extrapolating results to areas where data are limited.

In the previous Chapters we developed extensions to line transect methodology with which estimates of density and abundance may be obtained. As an example we applied the methods to estimate the relative abundance of ETP dolphin populations. In this Chapter we address issues related to the definition of ETP dolphin populations, and to the impact of such definitions on the resulting estimates of abundance. The benefits of monitoring ETP dolphin populations by area, rather than stock, given uncertainties in stock identification, are discussed. As an example we apply the method to stocks of offshore spotted dolphin. In addition, guidelines for a management strategy robust to uncertainty in stock structure are proposed.

4.2 Dolphin stocks in the ETP

Current management schemes assess the status of ETP dolphins on the basis of "stocks" defined as populations that are "substantially reproductively isolated" and "represent an important component of the evolutionary legacy of the species" (Dizon *et al.* 1994). Stock identification is based on a comparison of distributional patterns, demographic and physiological parameters, and phenotypic and genotypic traits between neighbouring populations (Dizon *et al.* 1992, 1994). Currently two stocks of spinner dolphins (eastern and white-belly) and two stocks of offshore spotted dolphins (northeastern and southern-western) are recognized (Dizon *et al.* 1994).

There are a number of ways in which the status of those ETP dolphin stocks is assessed (*cf.* DeMaster and Sisson 1992). Estimates of absolute abundance based on research vessel sightings surveys are available for a number of years, and trends in those estimates have been used to monitor dolphin stocks (Wade and Gerrodette 1992). In addition, estimates of absolute abundance for a given stock have been used to determine whether that stock was above its "maximum net productivity level", so that it could be considered to be at "an optimum sustainable population level" (e.g. Wade 1993, 1994). A population dynamics model, based on estimates of absolute and relative abundance, as well as estimates

of incidental mortality, has also been applied to determine the status of various stocks (Wade 1994). Finally, trends in annual estimates of relative abundance based on tuna vessel sightings data have been used since the 1980s to monitor ETP dolphin populations (Buckland and Anganuzzi 1988*b*, Anganuzzi and Buckland 1989, Buckland *et al.* 1992).

All of the methods mentioned above rely on the assumption that stocks have been correctly identified. When neighbouring stocks are easily distinguished in the field, such as eastern and whitebelly spinner dolphins, stock identification is greatly facilitated. The large morphological differences between eastern and whitebelly spinners suggest a high degree of reproductive isolation between the populations, which is the main criterion for defining ETP dolphin stocks. In contrast, when neighbouring stocks cannot be visually distinguished in the field, and they are not geographically isolated, the task of identifying stocks becomes more difficult; that is the case of the northeastern and southern-western stocks of offshore spotted dolphin. However, the intra-specific structure of ETP dolphin populations is characterized by a high degree of geographic variation. Take the case of spinner dolphins as an example. Even though eastern and whitebelly spinners are quite distinct morphologically, some individuals exhibit intermediate features of both stocks, suggesting the existence of a gradient, a cline, between the two stocks (Perrin 1990). In this case, the two forms are clearly mixing. How much mixing is required to occur before substantial reproductive isolation no longer exists? But if we cannot directly ascertain the rates of gene flow between the two stocks, how can we determine the degree of mixing that is taking place in the first place? In the case of offshore spotted dophins, these questions are even more difficult to address, since we can only hypothesize that a cline also exists between the two stocks, but we are not even able to determine that with any degree of certainty.

Stock identification in general, and in particular with respect to ETP dolphins, has long been recognized as a fundamental issue in the assessment and management of dolphin populations (e.g. Wade and Angliss 1997). Variation within ETP dolphin stocks is often discordant among the characters examined, suggesting that a number of stock divisions are possible (Tables 4.1 and 4.2, based on data taken from Barlow 1984, Barlow and Hohn 1984, Chivers and Myrick Jr. 1993, Dizon *et al.* 1991, 1992, 1994, Douglas *et al.* 1986, 1992, Hohn and Hammond 1985, Hohn *et al.* 1985, Perrin 1984, 1990, Perrin and Henderson 1984, Perrin and Reilly 1984, Perrin *et al.* 1979, 1991, 1994, Schnell *et al.* 1986). This led Dizon *et al.* (1992, 1994) to propose a system of stock classification that reflects this uncertainty. However, although the potential consequences of incorrect stock classifications have been examined (e.g. Wade and Angliss 1997), uncertainty in stock

identification has not been directly incorporated into the methodologies used to monitor the status of the stocks.

The methods used to assess the current status of ETP dolphin stocks also rely on the assumption that stock boundary locations have been correctly identified. Even if stocks have been correctly identified, the incorrect determination of stock boundaries can lead to erroneous inference about the status of the stocks. Take the northeastern stock of offshore spotted dolphins as an example, and let's first hypothesize that current boundary locations for that stock result in a range that is smaller than its true range, though they encompass the main area for that stock. If that stock is increasing, then one of two scenarios may take place: either there will be an increase in dolphin density within the assumed stock range or, if carrying capacity has been reached, then the range for that stock will expand. In the latter case, we would not be able to detect any trends in the population using current methodologies which rely on the correct placement of stock boundaries. Chances of incurring this type of error are greatest when stocks cannot be visually distinguished in the field. For example, while boundaries for the eastern and whitebelly stocks of spinner dolphin are asymmetric, and even overlap considerably, the northeastern and the southern-western stocks of offshore spotted dolphins are divided by straight lines, with no region of overlap between the two. Is that a reflection of difficulties in distinguishing between the two stocks of offshore spotted dolphin, or is the structure of offshore spotted dolphin stocks really that different from that of the stocks of spinner dolphins? As is the case with stock identification, uncertainty in the determination of boundary locations has not been directly incorporated into abundance estimation methods.

In addition, management of ETP dolphins on the basis of stocks relies on the existence of static, well-defined boundaries between stocks. However, the ETP is a dynamic environment, in which marked changes in oceanographic conditions take place as a result of El Niño events (Fiedler *et al.* 1992). These oceanographic changes affect both primary and secondary productivity, and coincide with shifts in the distribution of a variety of species. Given the observed association between dolphin distribution and oceanographic conditions in the ETP (e.g. Reilly 1990), it is likely that stock boundaries are dynamic, non-static. In strong El Niño years, for example, estimates of dolphin relative abundance invariably decline, but later increase (e.g. Lennert-Cody *et al.* In prep). In such years there is a drop in estimates of encounter rate and mean school size, and to a lesser extent in estimates of the effective strip (half-)width. It is possible that these smaller schools are more difficult to be detected, and also tend to be sighted closer to the vessels, which would explain the lower estimates of encounter rate and effective strip (half-)width. An

Hypothesis	Pro	Con
1 stock with incomplete mixing	- Regional patterning in morphology (animals farther apart exhibit greatest differences); - East/west or concentric cline in colour patterns and/or morphology *may* exist; - Peak in reproductive season overlaps between 0–5° N (but is this a clinal change or region of overlap?).	- Tagging data (albeit limited - 3% recoveries) indicate lack of extensive movement by N animals; no individuals crossed the Equator; - Hiatus in distribution; - Differences in the proportion of mature, pregnant and pregnant + lactating females between N and S forms (but there are problems with sampling biases); - S, W and N (NE) forms inhabit different habitats.
2 stocks (N (= NE + W), S) with no mixing	- Distance between W and S animals is great; - Differences in reproductive indices between NE (+ W?) and S animals exist; - No evidence of N animals moving south across the Equator; - Seasonal east-west movement by NE animals provides potential for mixing with W individuals.	- S and W animals morphologically more similar to each other than to NE individuals; - S and W habitats more similar to each other than to NE habitat.
2 stocks (NE, S+W) with no mixing	- No observed movement of NE animals south across the Equator; - Morphological similarities between S and W animals, which differ from NE individuals; - Different peak in calving season between NE and S individuals; - Habitat similarities between S and W animals, which differ from NE habitat.	- Distance between S and W animals is great (*cf.* limited movement of N animals; same pattern may be observed for S/W individuals); - East/west seasonal movement of NE individuals provides potential for mixing with W animals.
3 stocks (NE, S, W) with some degree of mixing (?)	- Morphological differences between S+W and NE animals, coupled with great distance between S and W individuals; - N animals never observed crossing the Equator (limited amount of movement).	- S and W animals similar morphologically; - S and W animals inhabiting similar habitat.

Table 4.1: Hypotheses about potential stock divisions for offshore spotted dolphins. NE = northeastern; S = southern; W = western.

Hypothesis	Pro	Con
1 stock with incomplete (local) mixing	- Eastern and whitebelly overlap in distribution (at least partially); - Extreme morphological differences at opposite ends of their range (radial pattern of variation); - Regional patterning in morphology (animals farther apart exhibit greatest differences); - Limited tagging data indicate eastern spinners have even more restricted movement pattern than N offshore spotted; - No detectable differences in mtDNA between whitebellies and eastern form, suggesting some degree of mixing.	- Differences in tooth morphology between eastern and whitebelly spinners, and between N and S whitebellies (but potentially attributable to different exploitation levels); - Differences in the proportion of pregnant + lactating females between eastern (higher rates) and N whitebellies; - Differences in morphological indices of maturity between male eastern spinners and male N whitebellies (but difference more likely to reflect different degree of exploitation).
2 stocks (eastern, N+S whitebelly) with some mixing	- Morphological differences between eastern and whitebelly forms; - Partial overlap in distribution; - Evidence for limited movement by eastern spinners; - Differences in age, growth and reproductive parameters between eastern spinner and N whitebelly (but may simply reflect different exploitation levels of the two forms).	- Lack of geographic concordant variation within whitebellies; - Great geographic distance between N and S whitebellies; - Differences in tooth morphology between N and S whitebellies; - No detectable differences in mtDNA between whitebellies and eastern form (but could be used as a pro).
3 stocks (eastern, N whitebelly, S whitebelly) with some mixing	- Hiatus (low density of animals) in whitebelly distribution around the Equator; - Differences in tooth morphology between eastern and whitebelly spinners, and between N and S whitebellies; - Evidence for limited movement by eastern spinners; - Differences in gross reproductive rates between eastern spinner, N and S whitebellies (but could be attributed to different levels of exploitation).	- Lack of geographic concordant variation within whitebellies; - No detectable differences in mtDNA between whitebellies and eastern form.

Table 4.2: Hypotheses about potential stock divisions for spinner dolphins. N = northern; S = southern.

alternative explanation would be that dolphin schools have broken up into smaller ones, dispersed over a larger area. However, a large number of schools would have to disperse beyond the stock boundary in order to have any substantial effect on the estimates of relative abundance; this is unlikely to be the case.

Since for management purposes ETP dolphin stocks are effectively treated as isolated units, it is also important that movement between stocks be negligible or non-existent. For example, we may fail to detect negative trends in abundance of a given stock, even though that stock is truly declining, if such decline is balanced by immigration from neighbouring populations. Again, as in the case of incorrectly determining boundary locations, chances of incurring this type of error are greater for stocks which are more difficult to be distinguished in the field, such as the northeastern and southern-western stocks of offshore spotted dolphin.

4.3 An alternative approach

In the last decade there has been a trend in conservation biology to incorporate uncertainty about a number of population characteristics or parameters into estimates of abundance, especially within a management context (e.g. Forney 2000). In the case of ETP dolphins, however, uncertainties surrounding the identification of ETP dolphin stocks, as well as the correct location of stock boundaries, have not been directly taken into account.

Perhaps the most obvious way of incorporating this uncertainty into assessment methods is by varying the location of stock boundaries and obtaining estimates of absolute or relative abundance under the various scenarios. However, this task is complicated by the fact that we are often dealing with a series of data, since the main interest is in the monitoring of trends. Hence, not only would we have to vary the boundary locations within each year, but we would also have to produce a number of combinations of varying boundary locations which would reflect variation between years.

As an alternative we propose that trends in dolphin abundance be monitored by small areas, rather than stock. That way we do not need to rely on the assumption that stocks have been correctly identified, and that boundary locations have been correctly determined. However, we must carefully choose both the size and the location of those small areas.

If the areas are too small, then trends in dolphin abundance within each area will not

be very informative, as they will likely exhibit considerable annual variability. However, they cannot be almost as large as the stock areas. Since we are particularly interested in trends near or around the northeastern offshore spotted dolphin stock boundary, we let the boundary lines correspond to the dividing line between neighbouring small areas. Given this predetermined location of the stock boundary lines, we chose to use $7° \times 10°$ latitude–longitude squares (Figure 4.1). This way the small areas have approximately the same size throughout the study region, and also we end up with a few small areas surrounding the boundary lines which separate the northeastern and the southern-western stocks of offshore spotted dolphin.

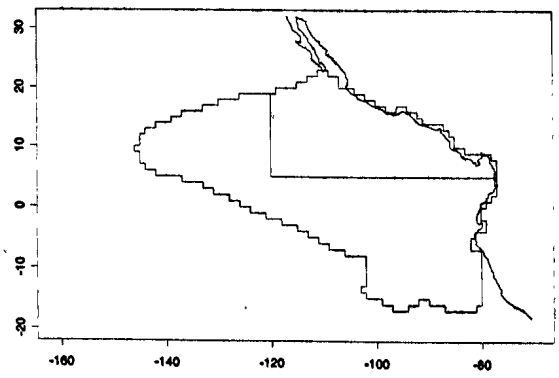


Figure 4.1: Location of small areas, determined with the northeastern and southern-western stock boundary as a starting point. Solid lines indicate stock boundaries for offshore spotted dolphins; dotted lines indicate boundaries of small areas.

We use the density surfaces obtained for offshore spotted dolphins from Chapter 2 to compute estimates of relative abundance for each small area for the years 1979–1997 (due to problems of data quality and coverage we omit the years 1975–1978; *cf.* Chapter 2). The density surface within each small area and each year is numerically integrated over the size of the small area, so that we end up with an annual series of point estimates of dolphin relative abundance for each small area. Trends in relative abundance within each small area are obtained by fitting a GAM to the annual series of point estimates, with 85% log-normal confidence intervals obtained based on the standard errors from the GAM model.

4.4 Results

Trends in offshore spotted dolphin relative abundance for each small area are shown in Figures 4.2 and 4.3. Although the large number of small areas makes the interpretation

of the results difficult, a few patterns can be observed. Trends in relative abundance near the coast appear relatively stable for both the northeastern and southern-western stocks of offshore spotted dolphin. In contrast, trends from small areas around the southern boundary between the northeastern and southern-western stocks exhibit considerable variability. Moreover, there appears to be a lag in the peaks in relative abundance between small areas north and south of the southern northeastern offshore spotted stock boundary, with the peaks in the southern areas occurring one to five years later than the ones in the northern areas. Not surprisingly, this lag is also observed between estimates of trends in relative abundance for the northeastern and southern-western stocks of offshore spotted dolphins (Figures 2.10 and 2.11). Nonetheless, no significant trends were observed in the majority of the small areas located around the stock boundaries. Trends in relative abundance over the southernmost region of the southern-western stock also appear relatively stable, although estimates of relative abundance for 1979 seem unduly large.

The location of small areas from Figure 4.1 was determined to a large extent based on current hypotheses about the structure of the stocks of offshore spotted dolphins. This is because the location of the boundary between the northeastern and the southern-western stocks was used as a starting point from which the small areas were generated. However, spatial patterns in trends in relative abundance over the different small areas may potentially vary depending on where the small areas are located. Say we have defined the small areas in a way such that we have at least one small area located over the southern boundary of the northeastern offshore spotted stock; that is, the small area encompasses areas both to the north and to the south of the stock boundary. If the southern boundary between the northeastern and the southern-western stocks of offshore spotted dolphin varies between years and say, for example, that individuals from the northeastern stock move south in certain years, and later return to the north, while the population of southern-western offshore spotted dolphins remains relatively stable over the period, then trends in relative abundance for small areas located over the southern boundary between the two stocks should be relatively stable. To investigate this, we generated a new grid of small areas in a way such that we ended up with a few areas located over the boundary between the northeastern and southern-western stocks of offshore spotted dolphin (Figure 4.4), and obtained estimates of trends in relative abundance for those areas as previously described (Section 4.3; Figures 4.5 and 4.6). However, the large number (25) of small areas made it difficult to discern any spatial patterns in the trends. Therefore we chose to pool the small areas, in a way such that we ended up with 8 small areas covering representative portions of the survey region (*i.e.* one in the westernmost area, three located over the northeastern

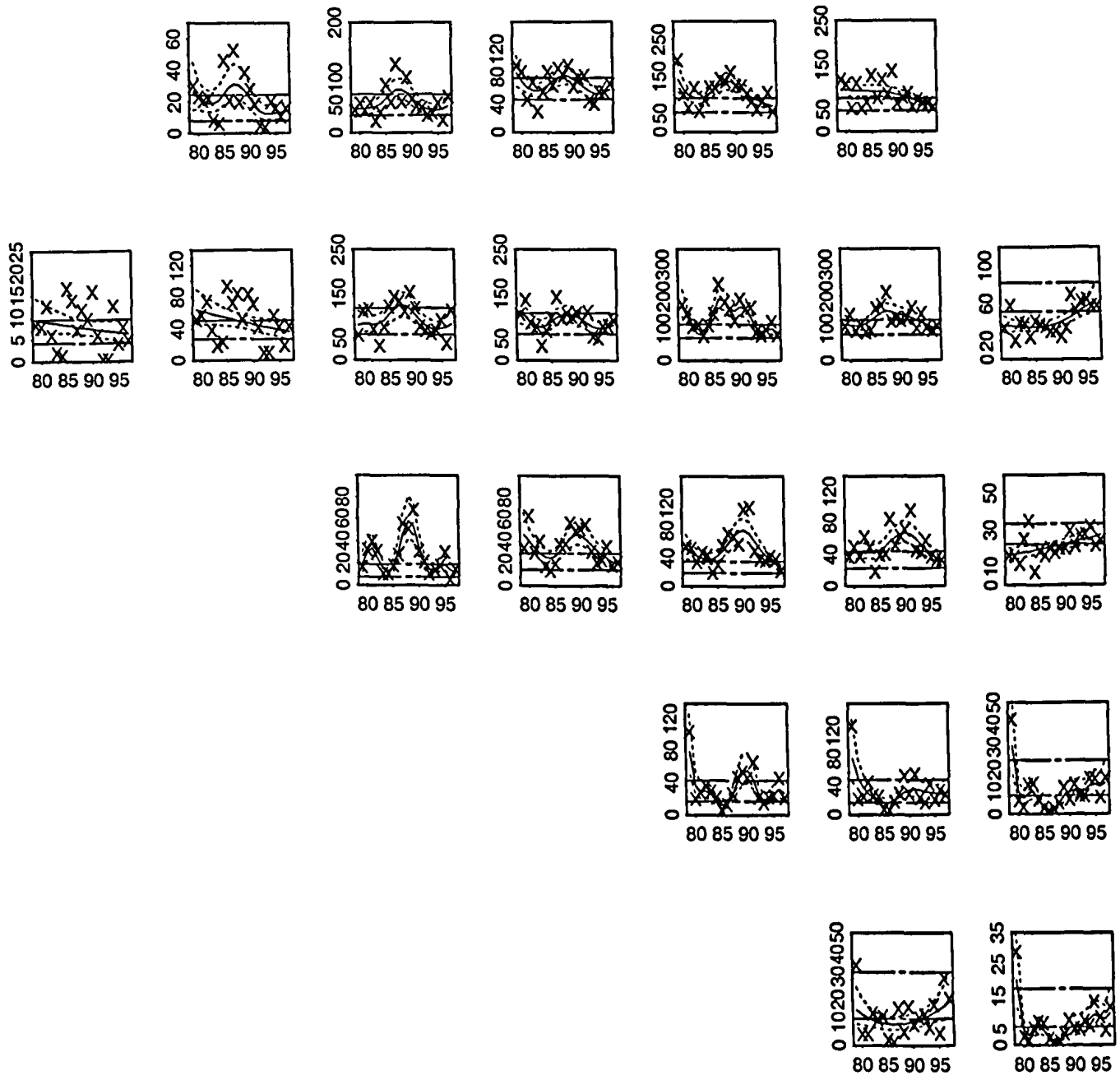


Figure 4.2: Trends in offshore spotted dolphin relative abundance by small areas, for 1979–1997. Each plot corresponds to each of the small areas shown in Figure 4.1.

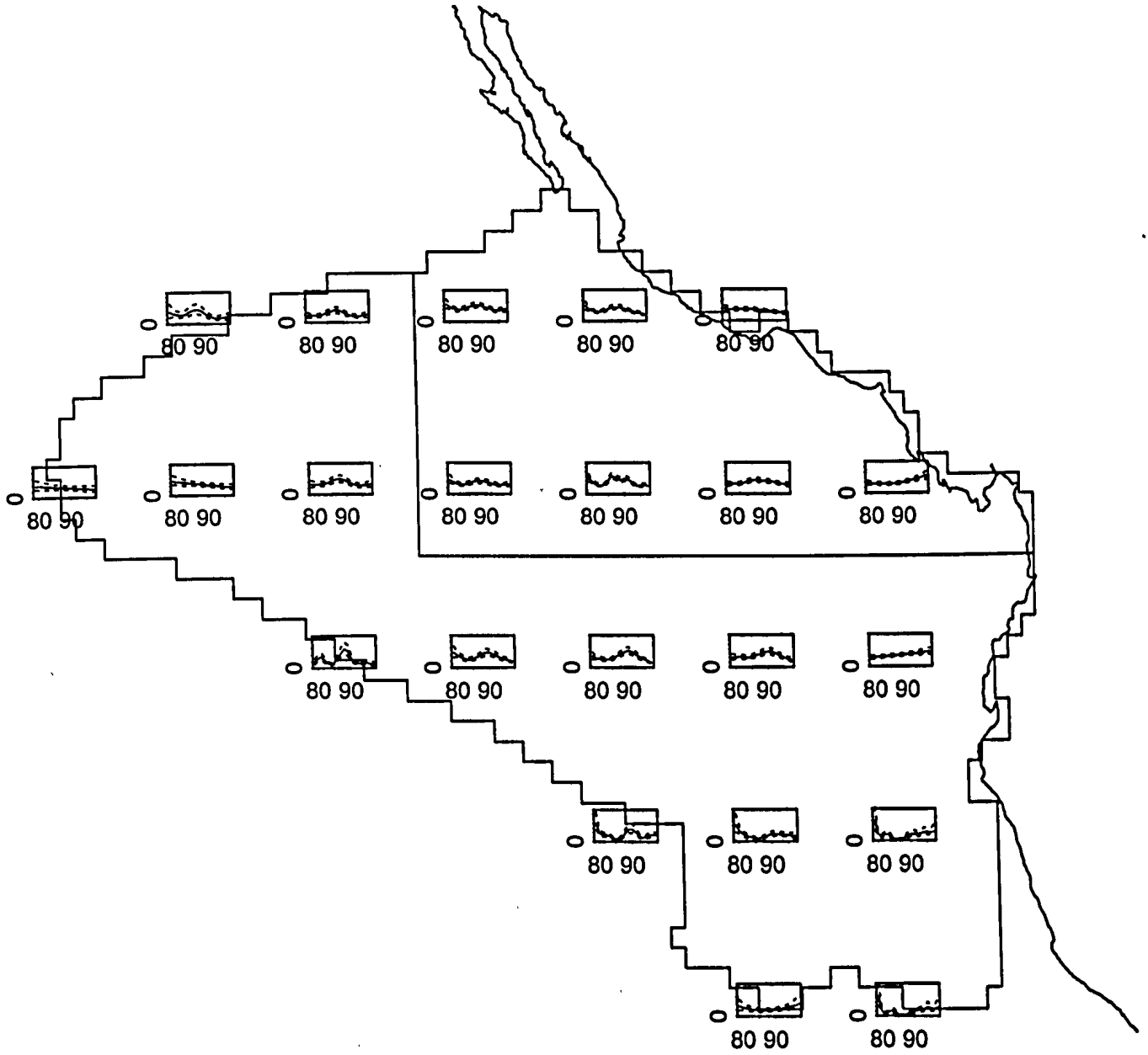


Figure 4.3: Trends in offshore spotted dolphin relative abundance by small areas, for 1979-1997, as in Figure 4.2. Each plot corresponds to each of the small areas shown in Figure 4.1. Solid lines indicate offshore spotted dolphin stock boundaries.

offshore spotted stock boundary, two completely within the northeastern stock area, and two over the southern area of the survey region; Figure 4.7).

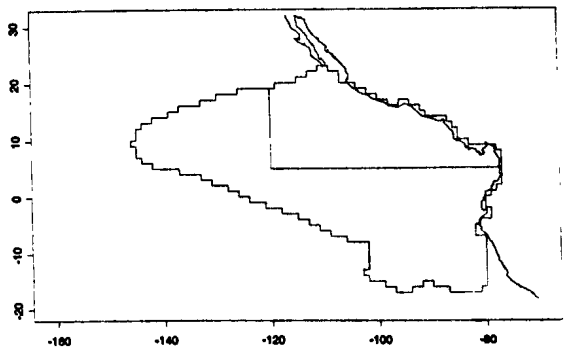


Figure 4.4: Location of small areas, with some areas placed over the boundary between the northeastern and southern-western stocks of offshore spotted dolphin. Solid lines indicate stock boundaries for offshore spotted dolphins; dotted lines indicate boundaries of small areas.

Trend estimates for the pooled small areas are presented in Figures 4.8 and 4.9. Trends in offshore spotted dolphin relative abundance appear quite stable near the coast, especially within the northeastern offshore spotted dolphin stock, and also within the southern area of the southern-western stock. Most of the fluctuation in relative abundance seems to occur around the western portion of the southern boundary of the northeastern offshore spotted stock, as well as within the central region of the northeastern stock and the western boundary of the northeastern offshore spotted stock. Nonetheless, in 5 out of the 8 areas no significant trends were observed when comparing relative abundance in 1997 with 1979. However, a significant decline since 1979 was observed for the small area within the central region of the offshore spotted dolphin stock, as well as for the one over the western region of the southern boundary for that stock. In contrast, a significant increase was observed near the coast, over the eastern portion of the southern boundary for the northeastern offshore spotted stock.

4.5 Discussion

In a review of the status of ETP dolphin stocks, DeMaster *et al.* (1992) point out that "geographically defined management units are not necessarily biologically meaningful", and also that "abundance can be estimated for a management unit, but trends in abundance must be determined by pooling stocks that are thought to mix or overlap in distribution". However, we have little information about the degree of mixing for some of the stocks,

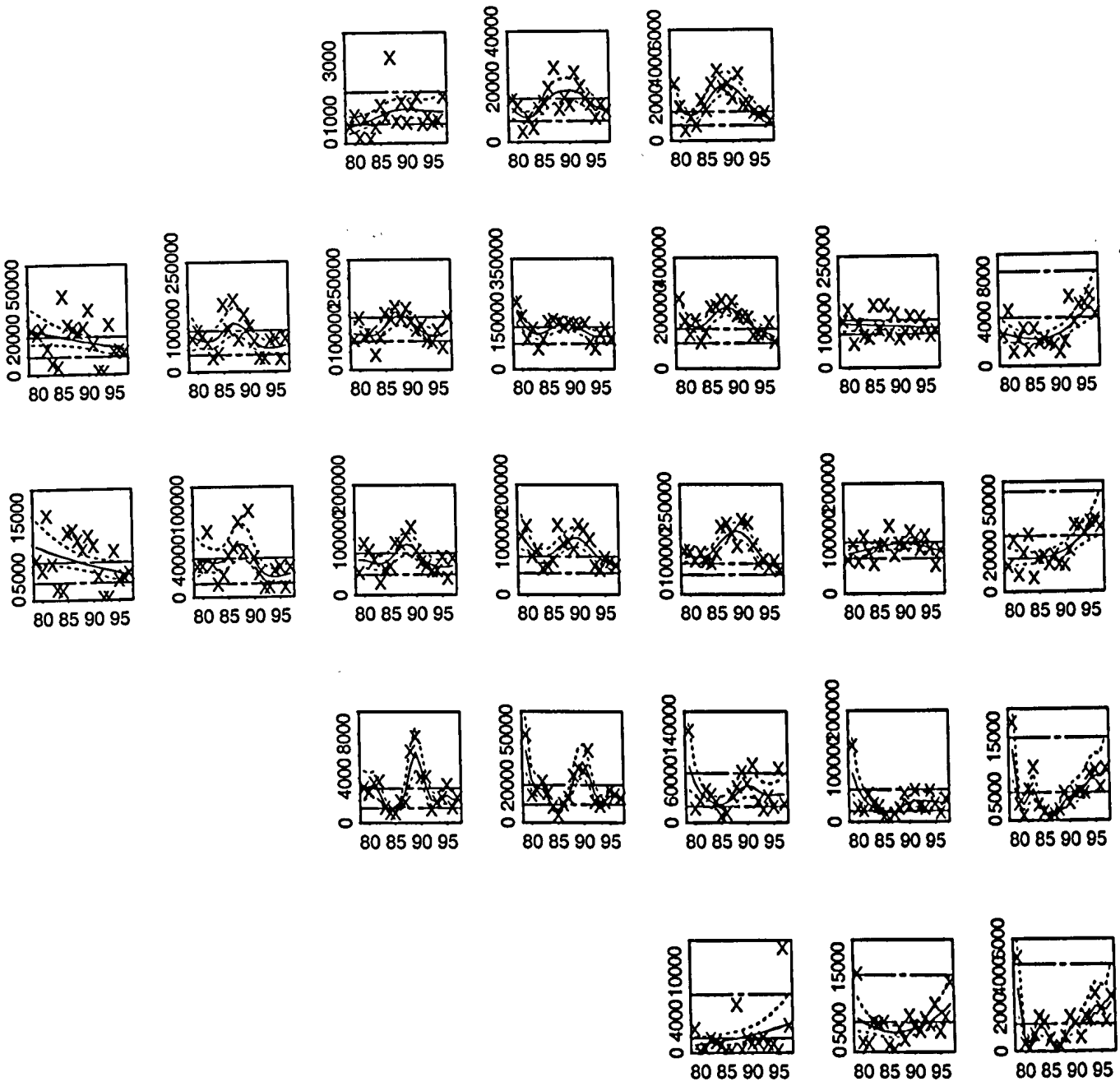


Figure 4.5: Trends in offshore spotted dolphin relative abundance by small areas, for 1979–1997. Each plot corresponds to each of the small areas shown in Figure 4.4.

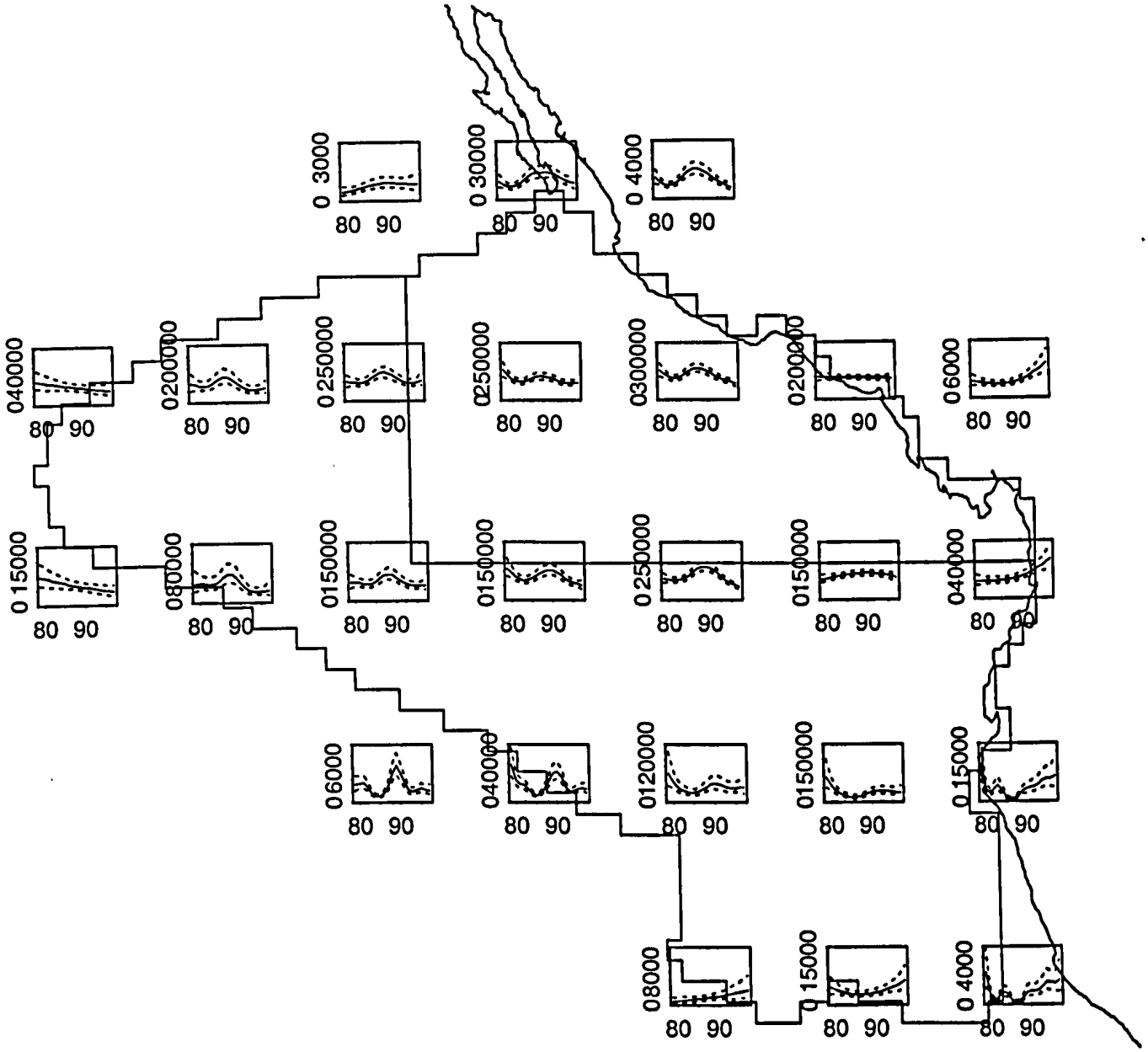


Figure 4.6: Trends in offshore spotted dolphin relative abundance by small areas, for 1979–1997, as in Figure 4.5. Each plot corresponds to each of the small areas shown in Figure 4.4. Solid lines indicate offshore spotted dolphin stock boundaries.

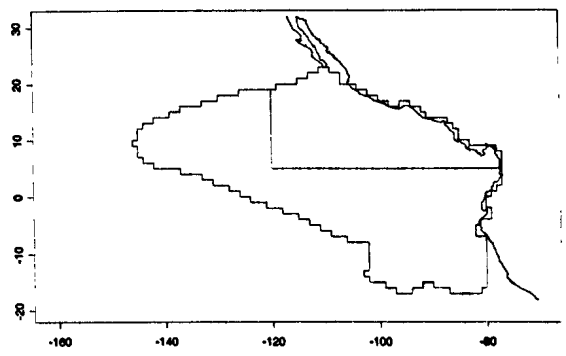


Figure 4.7: Location of pooled small areas. Solid lines indicate stock boundaries for offshore spotted dolphins; dotted lines indicate boundaries of small areas.

especially offshore spotted dolphins. Thus the monitoring of trends in dolphin relative abundance by small areas, rather than stocks, provides an alternative means of assessing the current status of ETP dolphin populations. Under this approach it is not necessary to define management units. In addition, any inference about the population of interest is no longer conditional on stock classifications being correct, or on any assumptions about stock structure.

The interpretation of the results, however, is not as straightforward as the simple evaluation of trends in abundance by stock, since we are no longer dealing with a single measure, but instead with a composite picture of the spatial variability in dolphin relative abundance over time. A significant decline in relative abundance since 1979 was observed within the northeastern offshore spotted stock area, as well as over the western portion of the southern boundary for that stock. Nonetheless, it is not clear whether or not these declines are directly linked (*i.e.* caused by movement of individuals) to the observed increase in relative abundance near the coast. Despite being highly variable, trends in relative abundance appear to be somewhat stable throughout the remainder of the range of ETP offshore spotted dolphins.

In addition to the assessment of the status of the populations, the observed spatial patterns in trends over small areas also provide information on the structure of ETP dolphin populations. Perhaps the most obvious result is the lack of concordance in spatial variation in the trends (*cf.* Figures 4.2, 4.5 and 4.8). One potential explanation is that, although the population is spread over a wide region, population responses occur at a much finer (local) spatial scale. It is also possible that this result is a reflection of spatial variability in environmental fluctuations, or of the degree of environmental variability across the re-

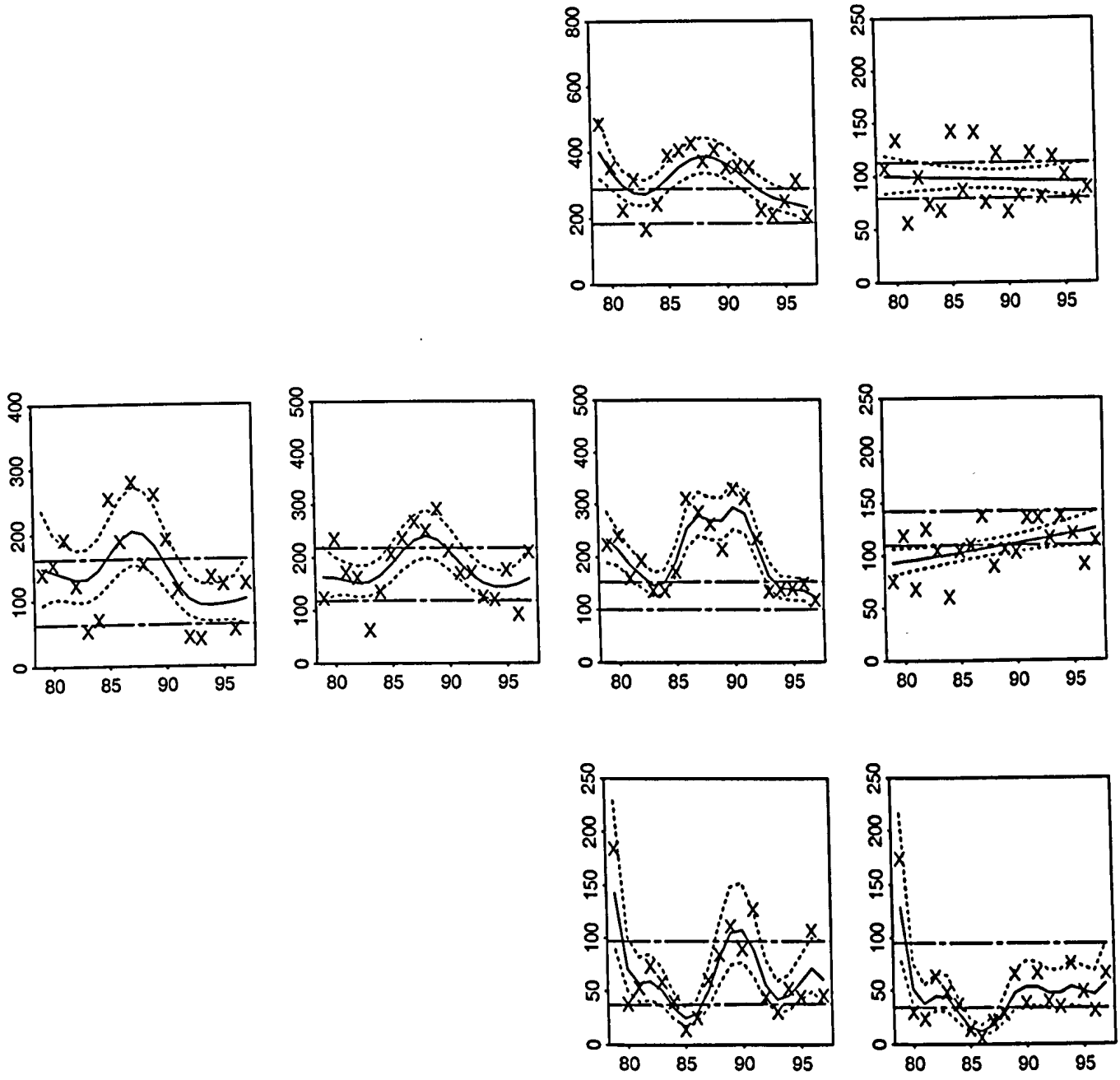


Figure 4.8: Trends in offshore spotted dolphin relative abundance by small areas, for 1979–1997. Each plot corresponds to each of the small areas shown in Figure 4.7.

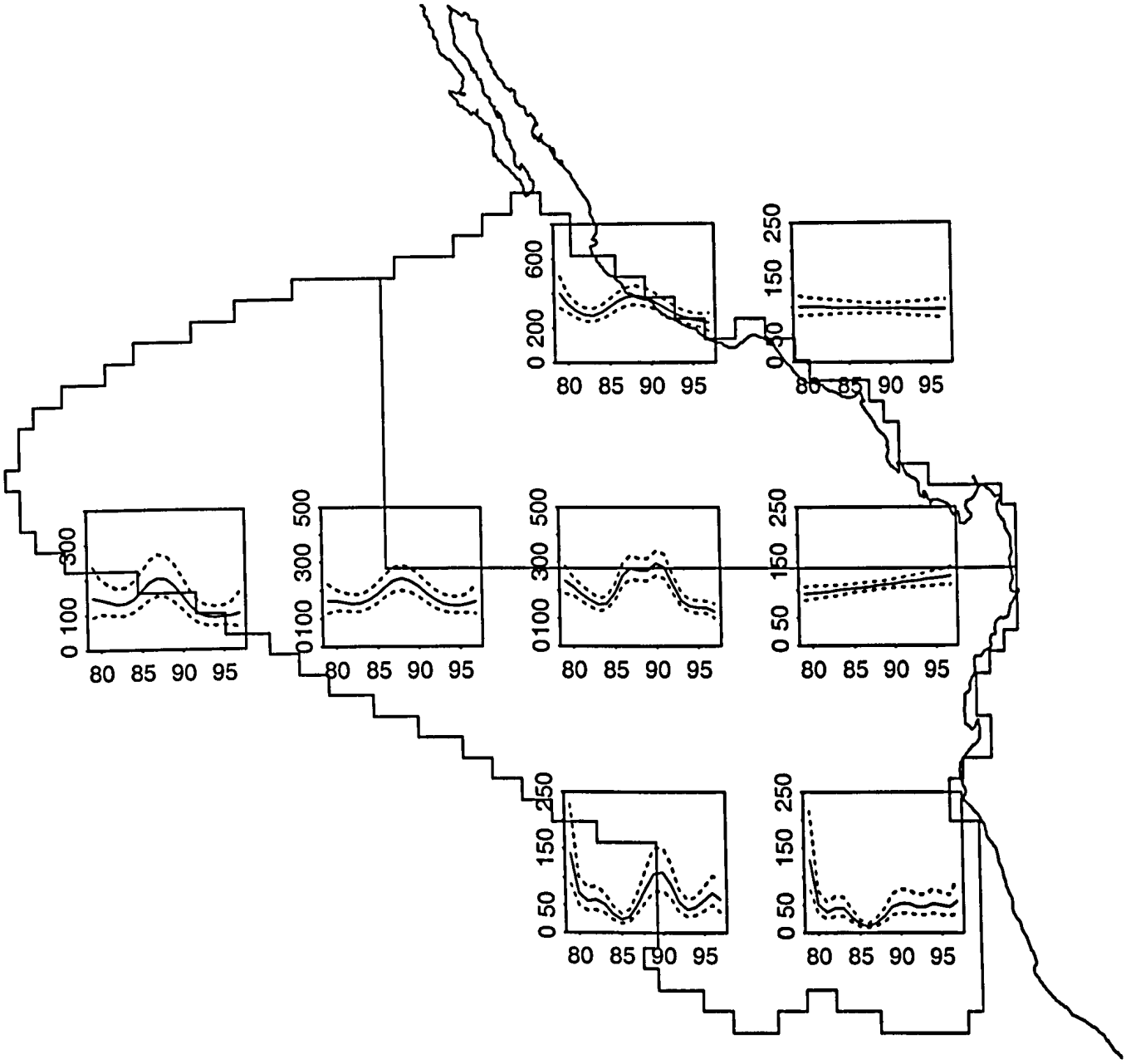


Figure 4.9: Trends in offshore spotted dolphin relative abundance by small areas, for 1979–1997, as in Figure 4.8. Each plot corresponds to each of the small areas shown in Figure 4.7. Solid lines indicate offshore spotted dolphin stock boundaries.

gion. In this case, the observed variability in trend estimates for certain small areas could then be caused either by movement of individuals, or by changes in the characteristics of the population (e.g. schools breaking up into smaller ones, and/or spreading over a wider area) which affect detectability. Regardless of the cause for this lack of spatial concordance in the trends, it is likely that this additional source of variability will reduce the precision of trend estimates, which in turn would decrease our ability to detect significant changes in the status of the populations. This effect is likely greater when trends are estimated on the basis of stocks.

The difficulty of interpretation of trend estimates by small areas precludes the complete replacement of the stock-based assessment approach by one based on small areas. On the other hand, the spatial variability in the trends reduces the efficiency of assessment methods based on stocks. In addition, some of this variability may be a result of movement of individuals across stock boundaries, which exemplifies the need to incorporate uncertainty about stock structure into assessment methods. Clearly both methods have limitations, but they provide complementary information. Therefore, their combined used as a means of assessing the status of populations is recommended.

Our results clearly indicate that both the size and location of the small areas affect the spatial patterns observed in the trends. In the case of offshore spotted dolphins, we started by generating small areas based on current hypotheses about the structure of that population, so that assumptions related to those hypotheses could be investigated. In addition, we also examined trends based on a grid of small areas which did not follow the stock divisions currently in use. By examining results for the two scenarios, we were less susceptible to biases resulting from the size and placement of the small areas that we chose to adopt. However, simulations are required to quantify how robust the results are to these variations in the size and location of the small areas, and also to determine whether there exists a combination of size and location of small areas which minimises spatial variability, so that the status of the populations can be more easily (and objectively) determined.

Chapter 5

General discussion

5.1 Introduction

The work presented in this thesis aimed at the development of methodology robust to violations of certain assumptions of line transect sampling. More specifically, we focused on alternative means of estimating abundance using data from platforms of opportunity, in which search effort is not random, and also of directly incorporating heterogeneity in detection probabilities into estimation procedures for $f(0)$. In addition, an alternative means of monitoring the status of populations, based on the examination of trends in abundance over small areas, rather than based on stocks, was also examined.

Each of those methodological developments touched a number of statistical issues (e.g. model specification, model selection, diagnostics, variance estimation, etc.), and due to time and computational constraints not all of those issues could be fully resolved. In addition, the tuna vessel observer data used in all examples presented a number of additional difficulties which resulted from the nature of the data set itself. Here we discuss possible avenues for future research in light of the results we obtained.

5.2 Spatio-temporal modelling

Model selection in the context of GAMs has been an active area of research in the last two decades or so (*cf.* Hastie and Tibshirani 1990). Despite being perhaps the most widely used model selection criterion, AIC tends to over-fit the models; that is, it tends to select more complicated models over simpler ones (Hastie and Tibshirani 1990). As an alternative, cross-validation and generalised cross-validation (GCV) have been proposed.

The general idea behind GCV is analogous to the jackknife procedure commonly used to estimate variances (*cf.* Efron and Tibshirani 1993). In short, the model is re-fitted n times (where n corresponds to the total number of observations), but at each one of those times one of the observations is omitted. Some measure of model fit (e.g. sum of squares) can then be computed based on the n fitted models. Until recently, calculation of GCV required considerable computing power. However, a new procedure proposed by Wood (2000) made GCV a more feasible tool to be used. Nonetheless, the large number of potential combinations of covariates to be included in the models, together with the various possible degrees of smoothing for each covariate, and the potential large number of observations, continue to make model selection based on GCV a very computer-intensive exercise.

The results from Chapter 2 clearly show that automated model selection can be misleading. For example, in the models for the number of schools, automated model selection using AIC tended to select relatively large (4 and 8) degrees of freedom for the smoothed covariates, yet examination of the predicted values from the models showed spurious results. Therefore, while perhaps better measures of model fit (such as GCV) may be employed to select between competing models, subjective decisions based on knowledge of the processes being modelled should not be completely discarded.

GAMs allow a suite of diagnostic tools to be used, ranging from standard and partial residual plots to more sophisticated graphics methods (e.g. plots of Cooke's distance, qq-plots). In the case of the tuna vessel observer data, however, the use of the more standard methods is not practical as the large number of observations makes it difficult to detect any patterns in the plots. We used qq-plots as diagnostic tools, since the simple relationship shown in the plots facilitate the visualization of any patterns even when the number of observations is large. Nonetheless, the use of qq-plots can be expanded by the computation of 'envelopes', so that the plots can become another tool (in addition to model selection criteria such as AIC or GCV) with which to test between competing models. To construct these envelopes, the procedure for computing the qq-plots can be repeated for models fitted to bootstrap resamples, from which 'percentile' confidence intervals for the qq-plots can be obtained. We could then test whether the original qq-plot fell within the envelope.

In the spatio-temporal models from Chapter 2, any autocorrelation between neighbouring cells was ignored; that is, we effectively assumed that spatial and temporal trends in the response variable resulted from covariate effects. Work by Augustin (1999) has shown

that abundance estimation based on the numerical integration of the predicted surface from spatio-temporal models is robust to the presence of autocorrelation. However, it is important that no spatial or temporal correlations be present in the residuals, as that may be an indication of a poor model fit. Plots of residuals against ordered values of latitude, longitude and month for the binomial model, the count model and the mean school size model from Chapter 2 for 1975–1997 are shown in Figures 5.1 through 5.9. There does not seem to be any strong spatial patterns in any of the models, as indicated by the plots of residuals against latitude and longitude. Some of the plots, however, clearly show the lack of observations over some of the spatial range. In contrast, plots of residuals against month appear to exhibit trends in some of the years. Some of those trends may be partially explained by lack of observations over portions of the survey region during certain months of the year. For example, effort near the western boundary of the survey region, between 120° and 146°W, is usually scarce during the winter months. However, some of those trends may also be due to autocorrelation, or to unmodelled variability. Distinguishing between autocorrelation and trends due to unmodelled variability is not, however, an easy task, and the need for further work in this area has already been highlighted (Augustin 1999).

5.3 Incorporating covariates into $f(0)$ estimation

Simulation results from Chapter 3 suggest that the proposed approach which directly incorporates covariates into $f(0)$ estimation is both accurate (unbiased) and precise, at least when only one covariate is included in the model. In the case where multiple covariates are used, simulation results showed the importance of carrying out covariate selection in order to obtain accurate and precise estimates of density.

As was the case with the spatio-temporal modelling approach from Chapter 2, we used qq-plots as a tool for model diagnostics. The use of 'envelopes', as described in the previous section, could also provide a more objective means of testing the model fit.

Borchers (1996) pointed out that observed variation in detection probabilities can be greater than analytic estimates based on estimators which include a single covariate. He argues that this may be a result of the effect of additional covariates which have not been included in the model. Therefore, it is recommended that robust estimators of the variance be used in place of analytic ones.

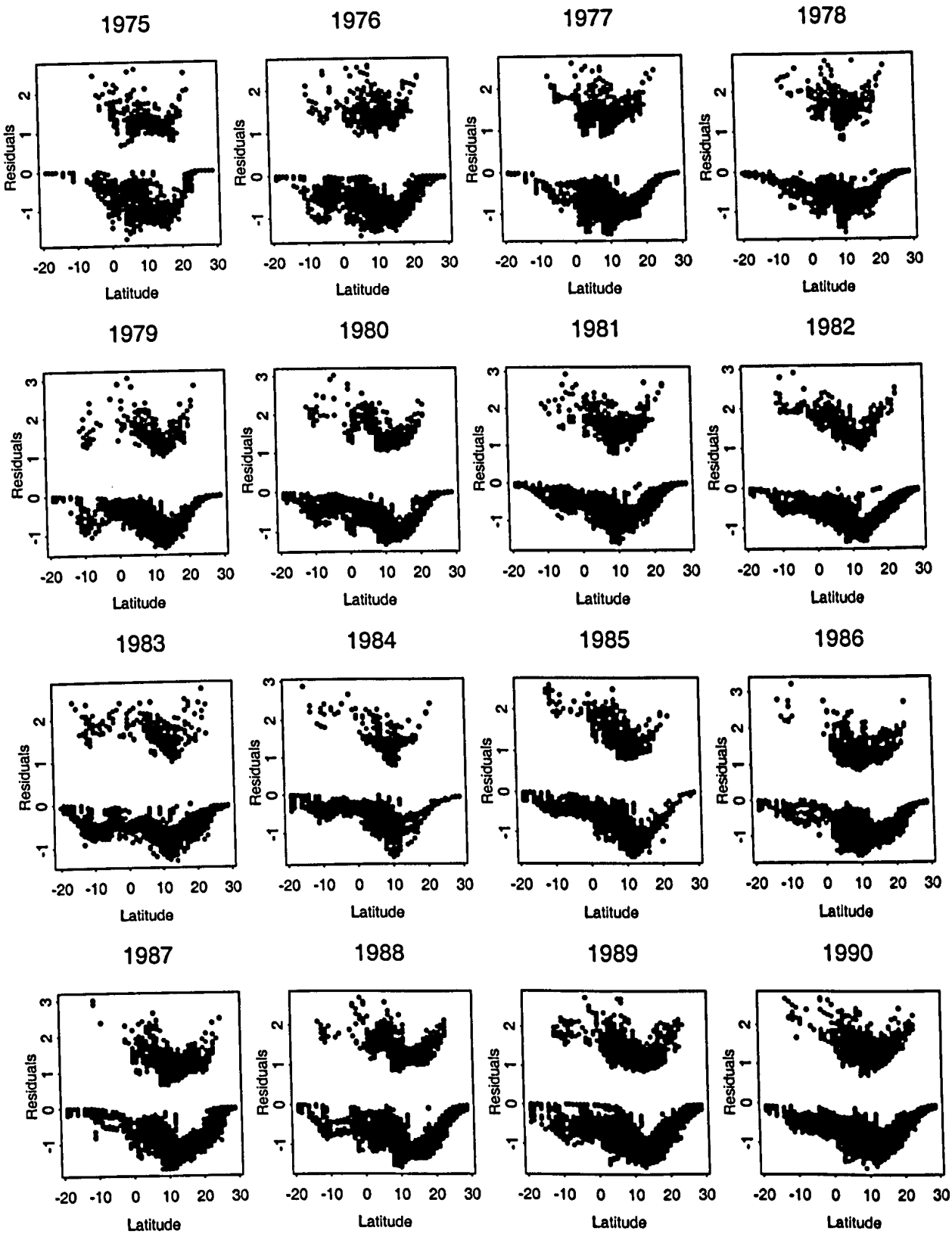


Figure 5.1: Plots of residuals against ordered values of latitude for the binomial model from Chapter 2, for 1975–1997.

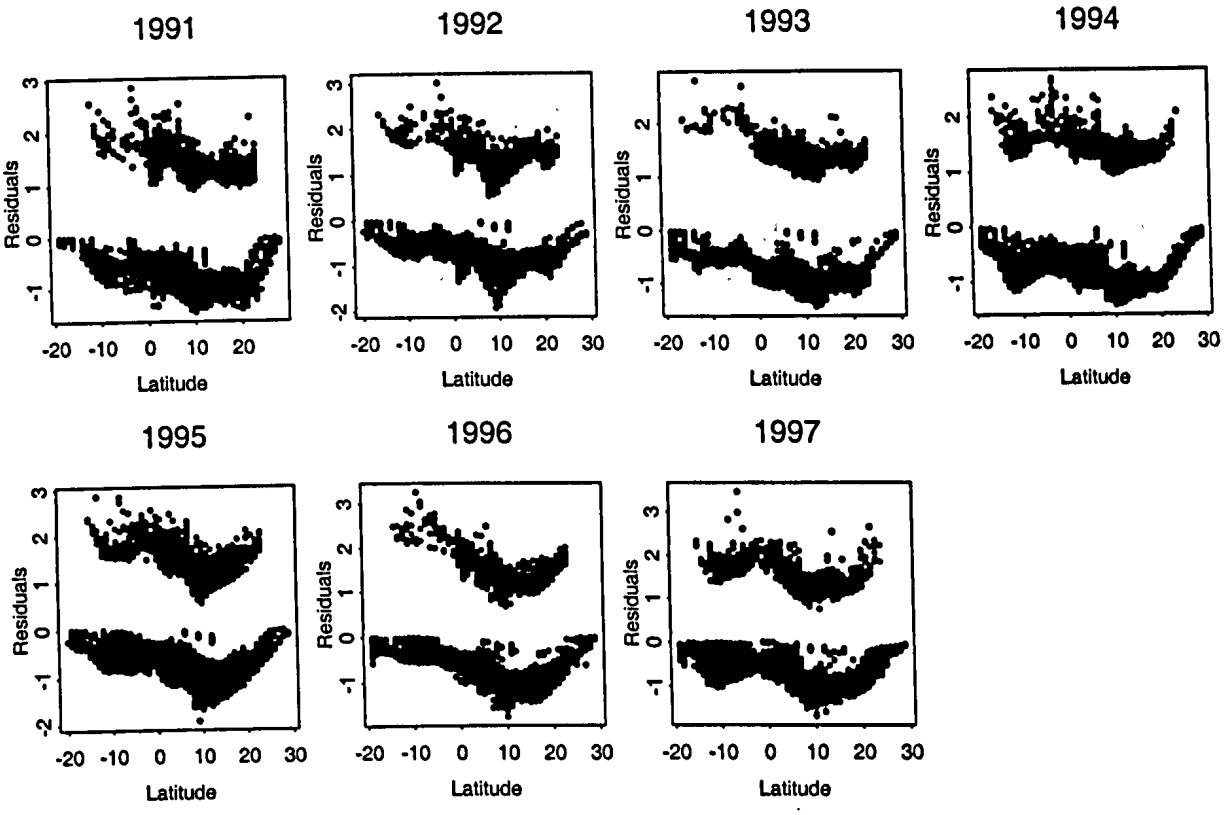


Figure 5.1: Plots of residuals against ordered values of latitude for the binomial model from Chapter 2, for 1975–1997.[continued]

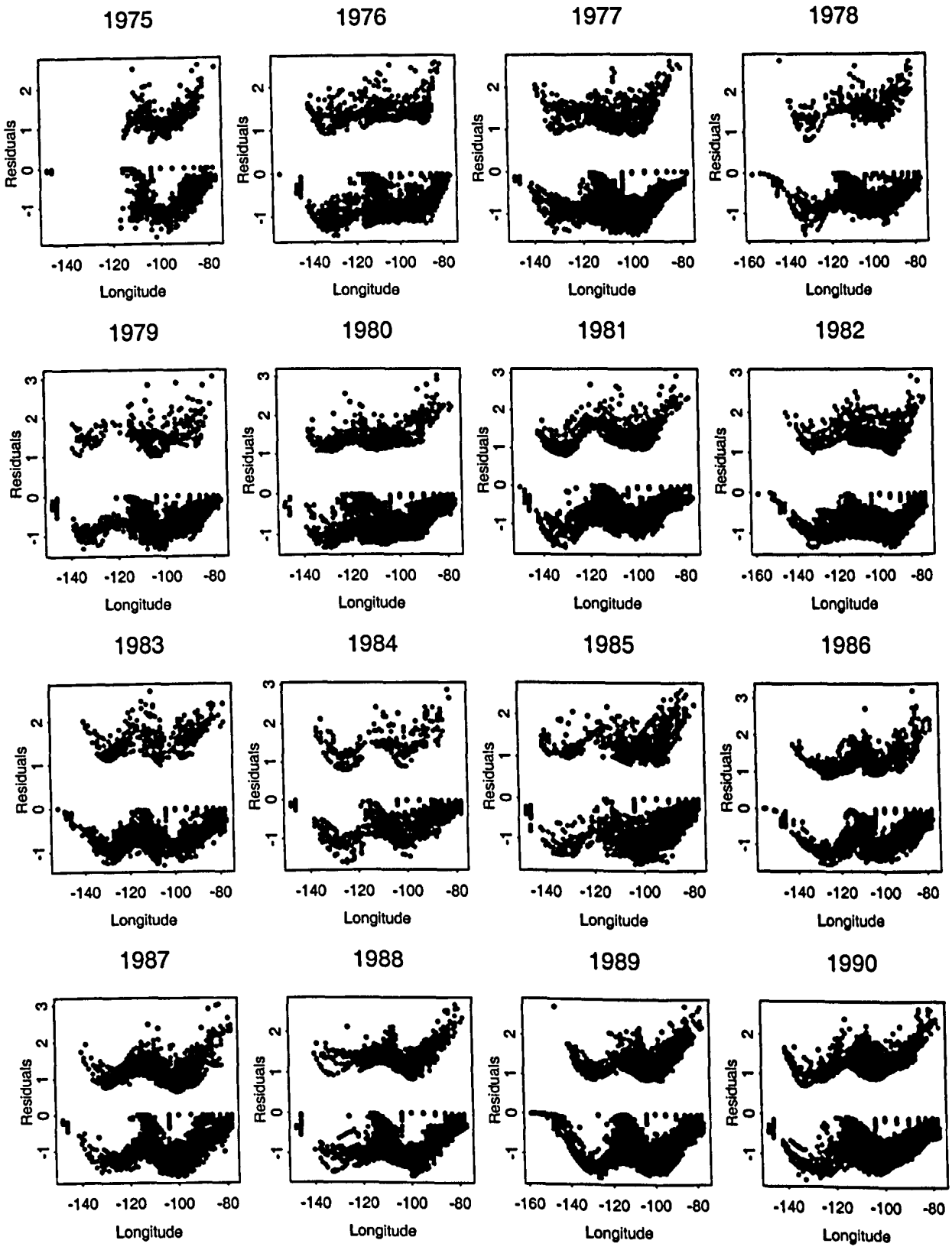


Figure 5.2: Plots of residuals against ordered values of longitude for the binomial model from Chapter 2, for 1975–1997.

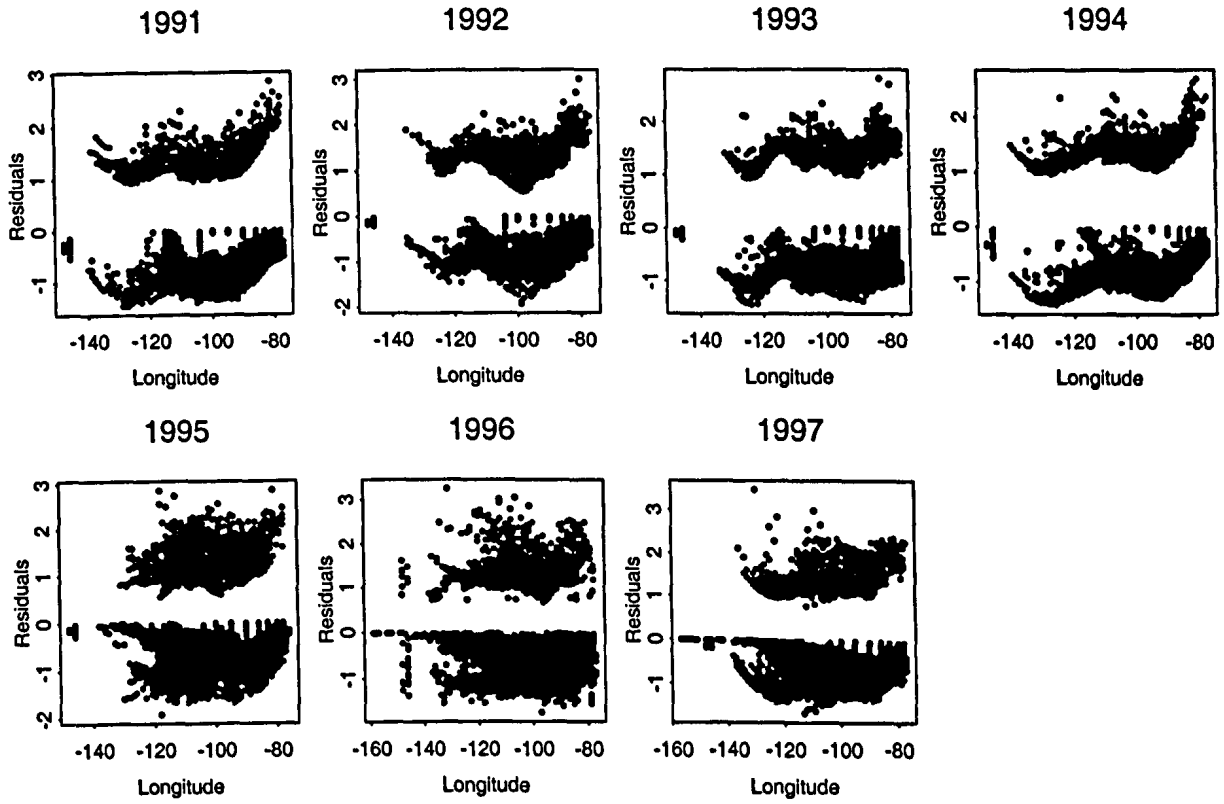


Figure 5.2: Plots of residuals against ordered values of longitude for the binomial model from Chapter 2, for 1975–1997.[continued]

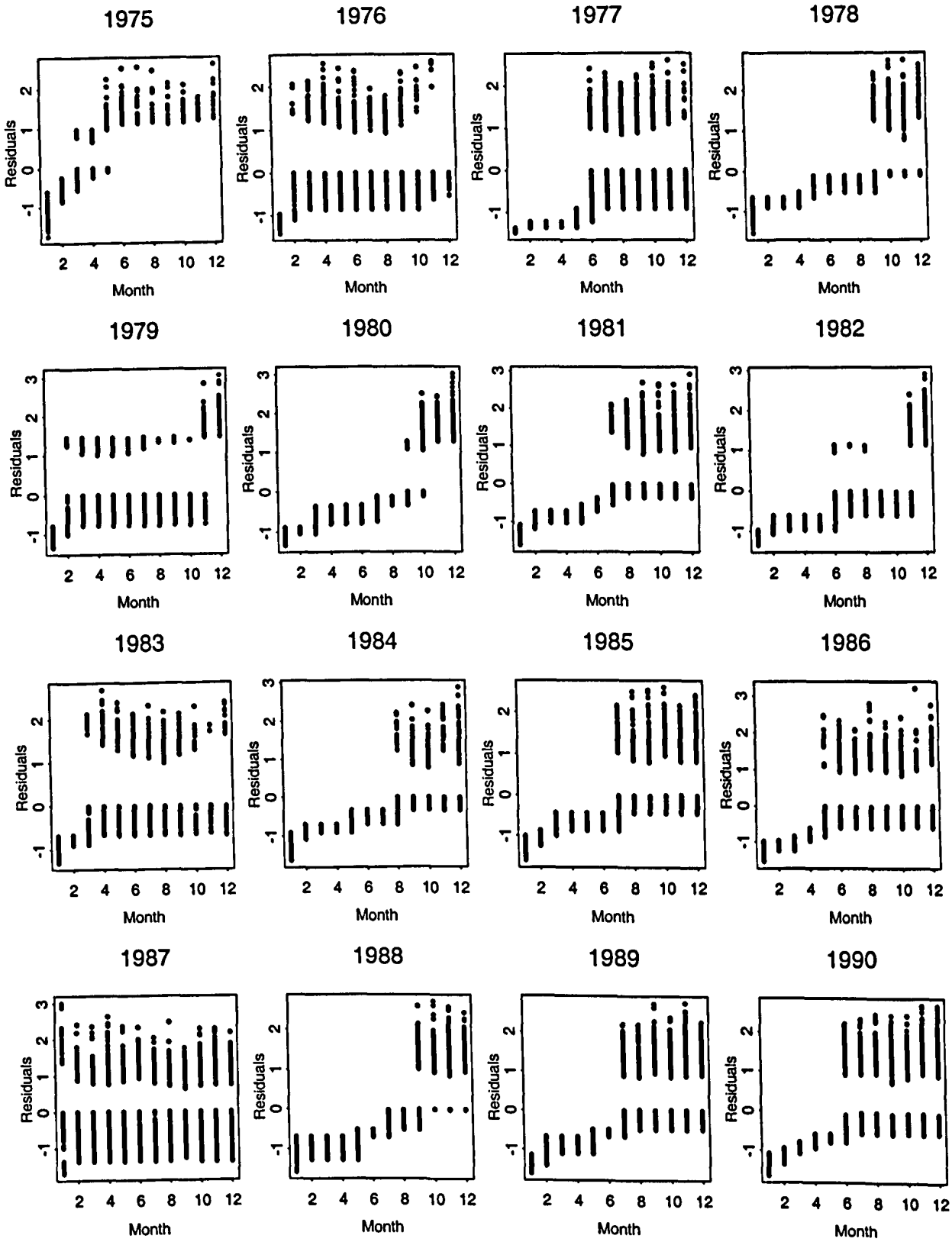


Figure 5.3: Plots of residuals against ordered values of month for the binomial model from Chapter 2, for 1975–1997.

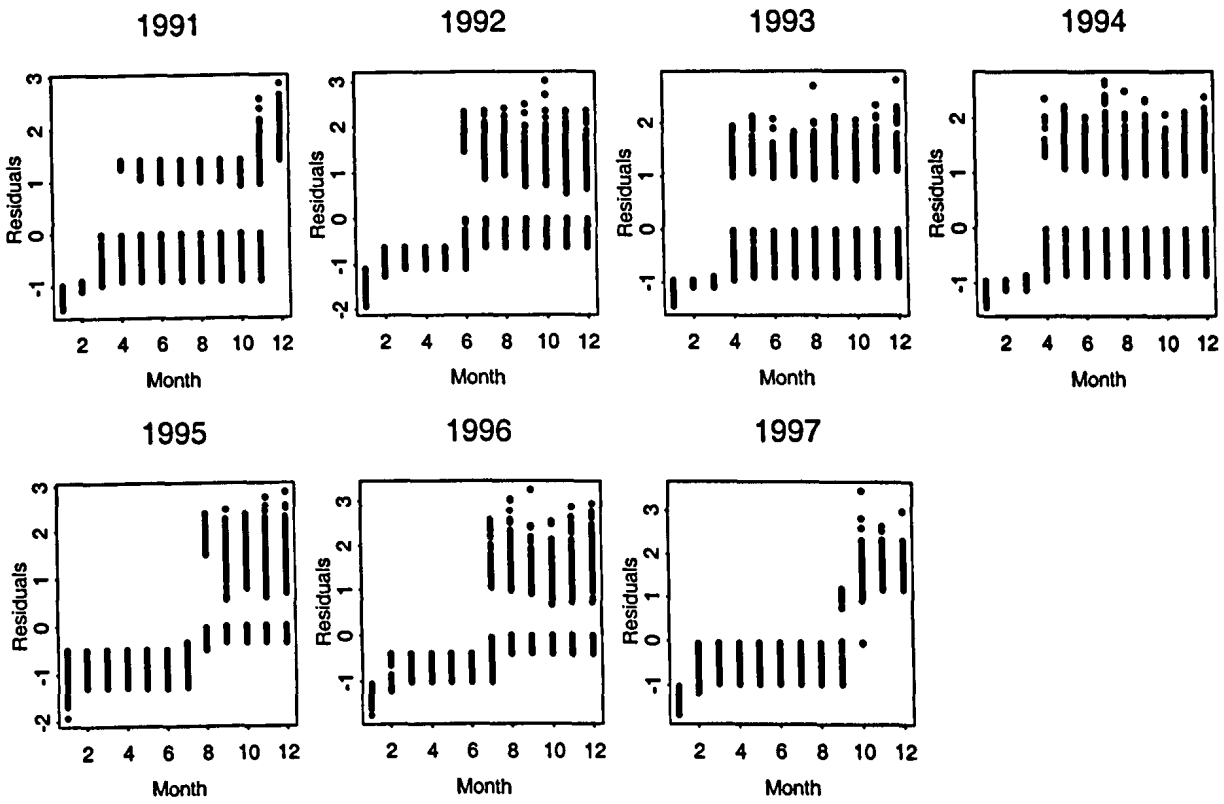


Figure 5.3: Plots of residuals against ordered values of month for the binomial model from Chapter 2, for 1975–1997.[continued]

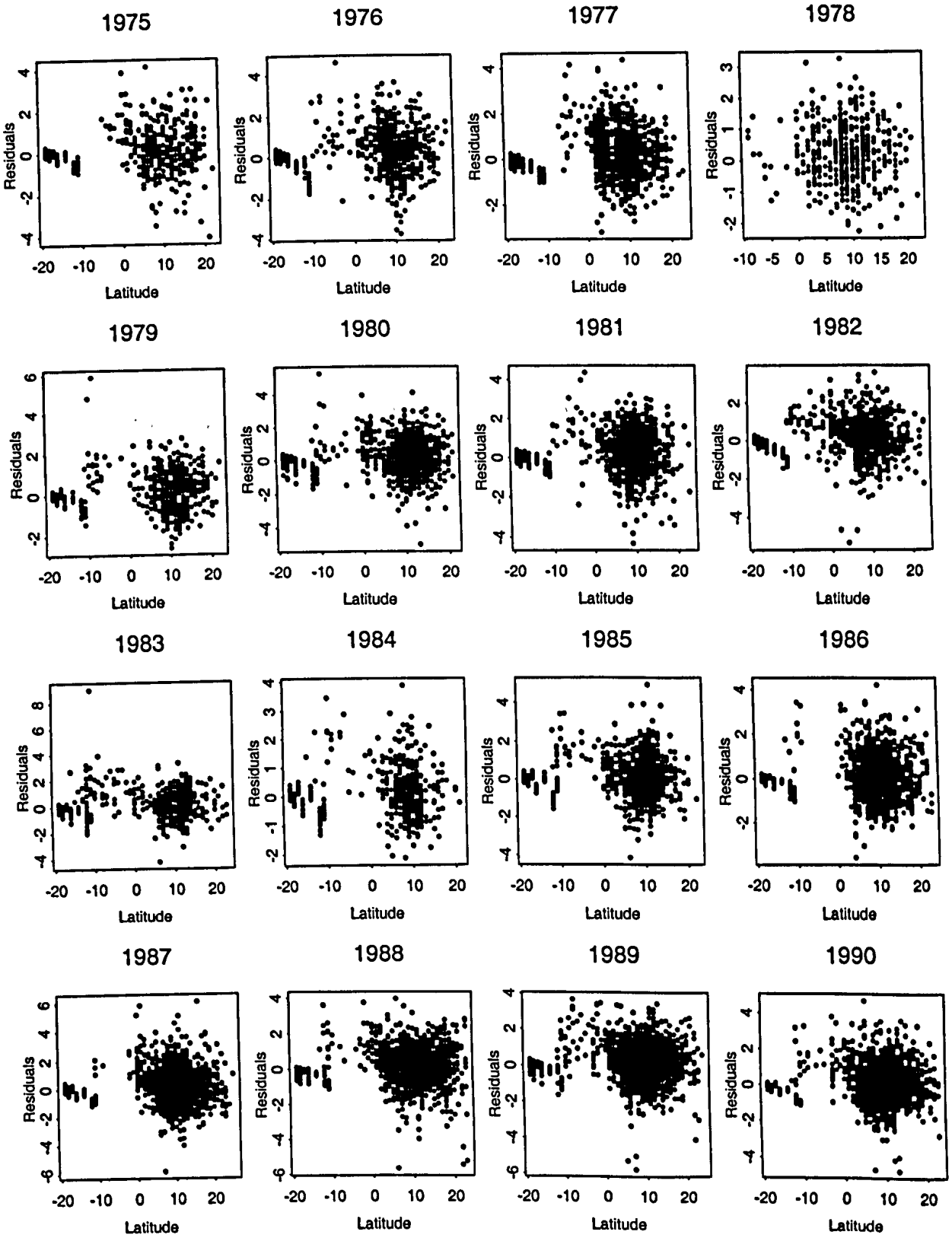


Figure 5.4: Plots of residuals against ordered values of latitude for the count model from Chapter 2, for 1975–1997.

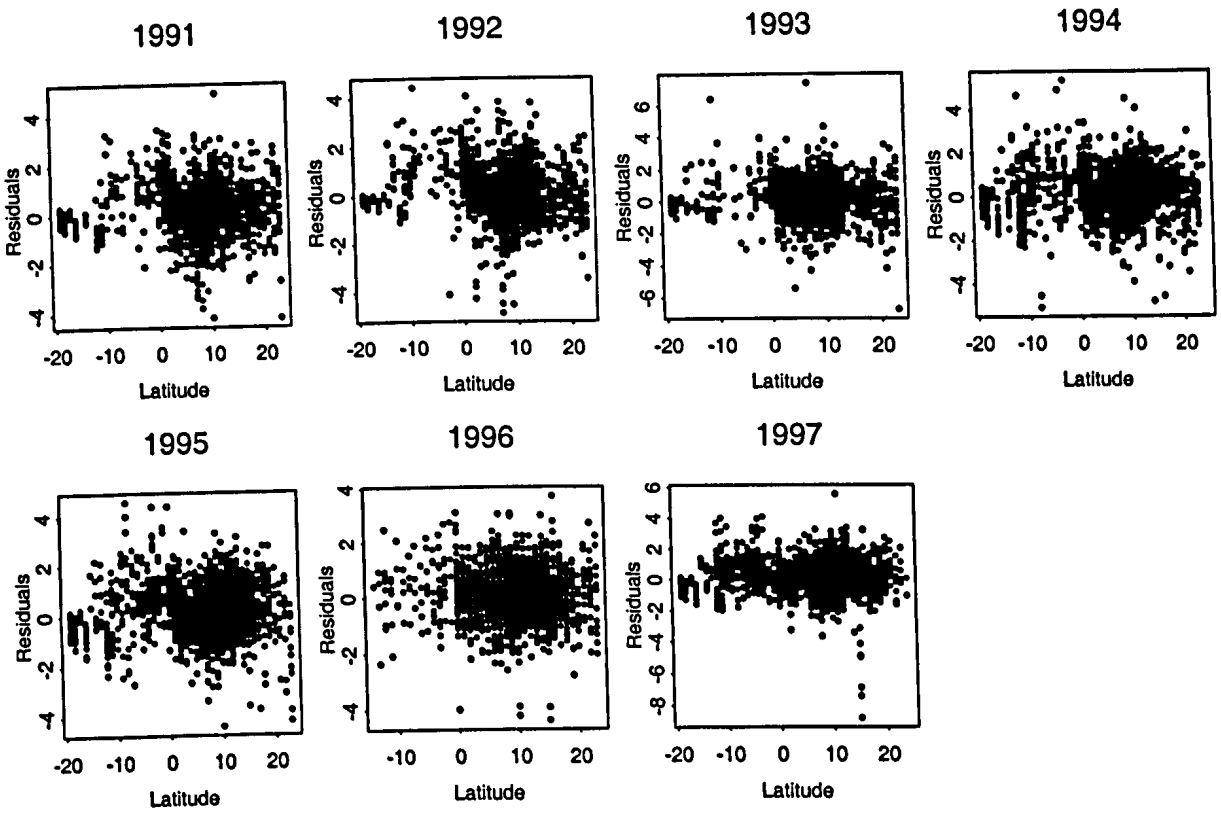


Figure 5.4: Plots of residuals against ordered values of latitude for the count model from Chapter 2, for 1975–1997.[continued]

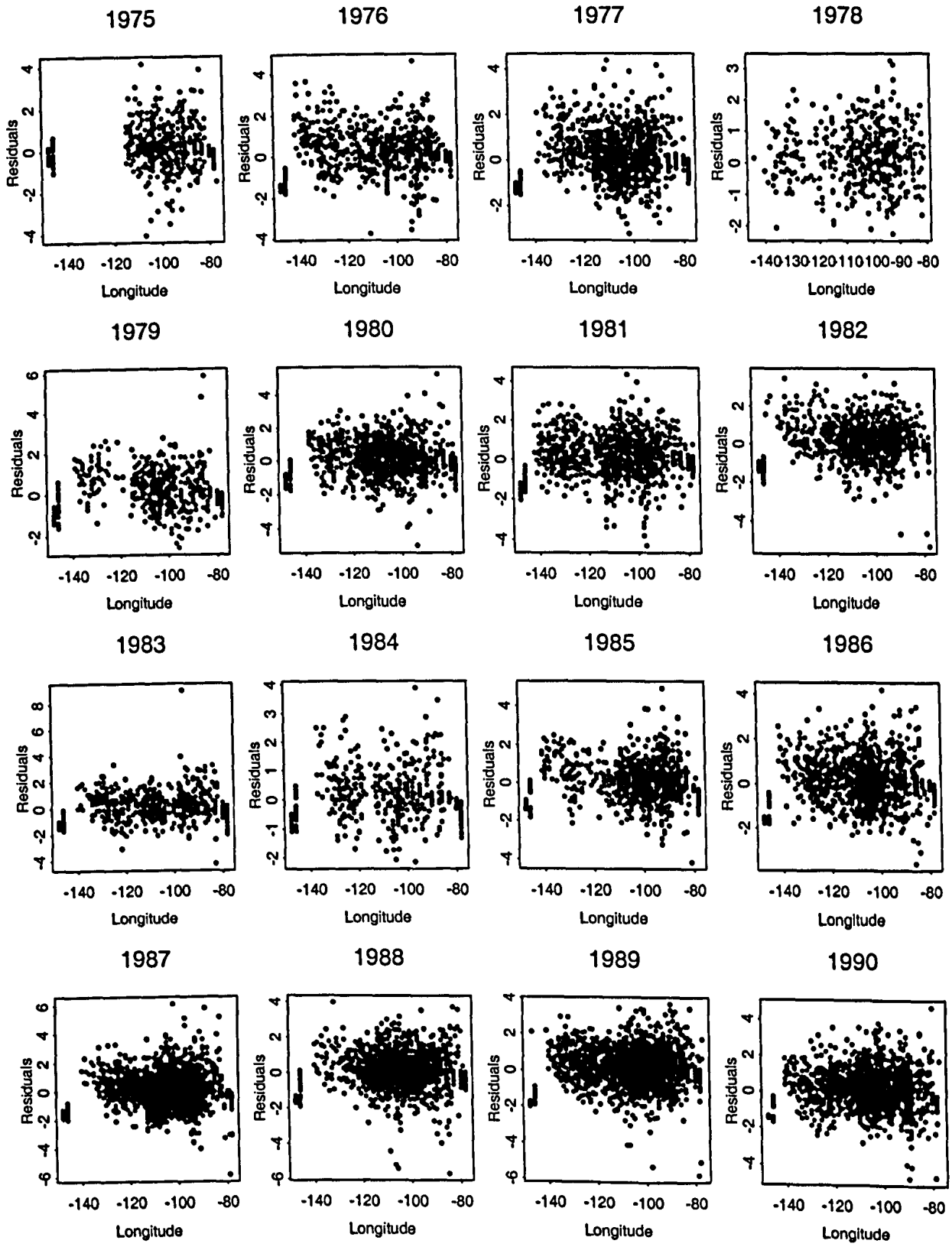


Figure 5.5: Plots of residuals against ordered values of longitude for the count model from Chapter 2, for 1975–1997.

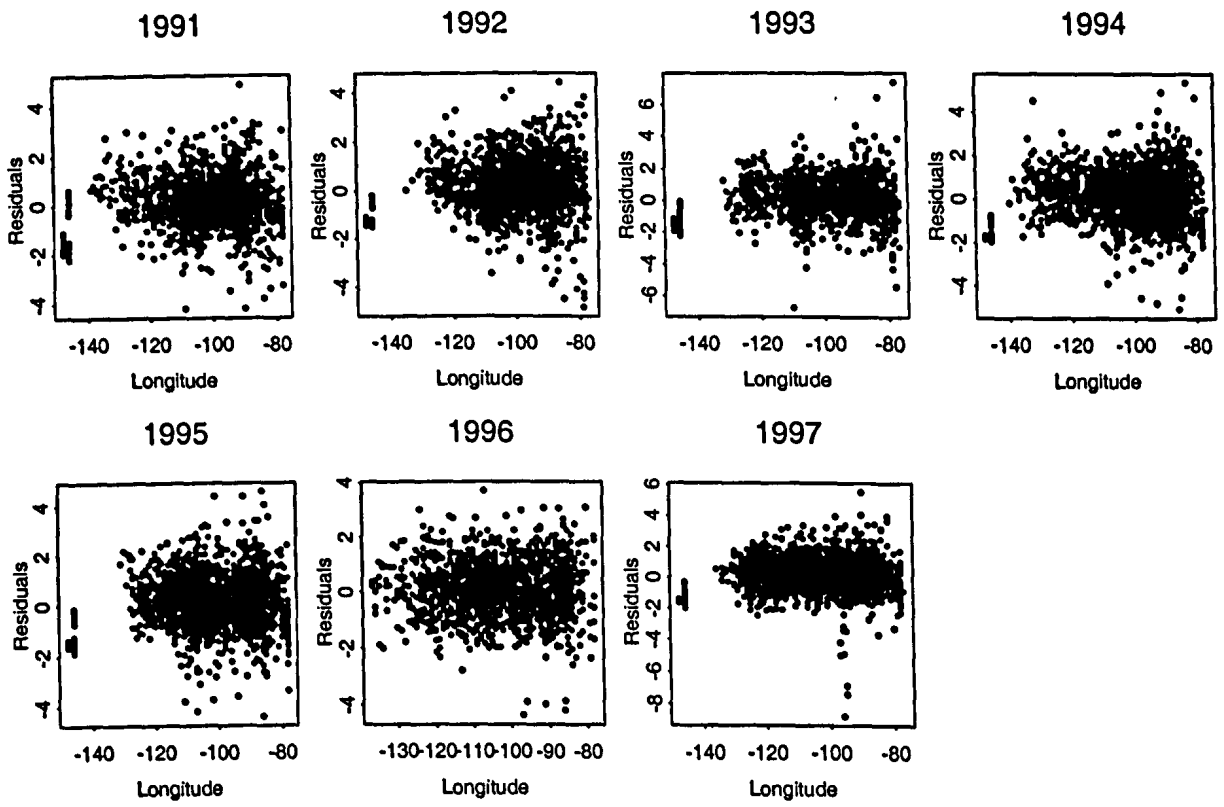


Figure 5.5: Plots of residuals against ordered values of longitude for the count model from Chapter 2, for 1975–1997.[continued]

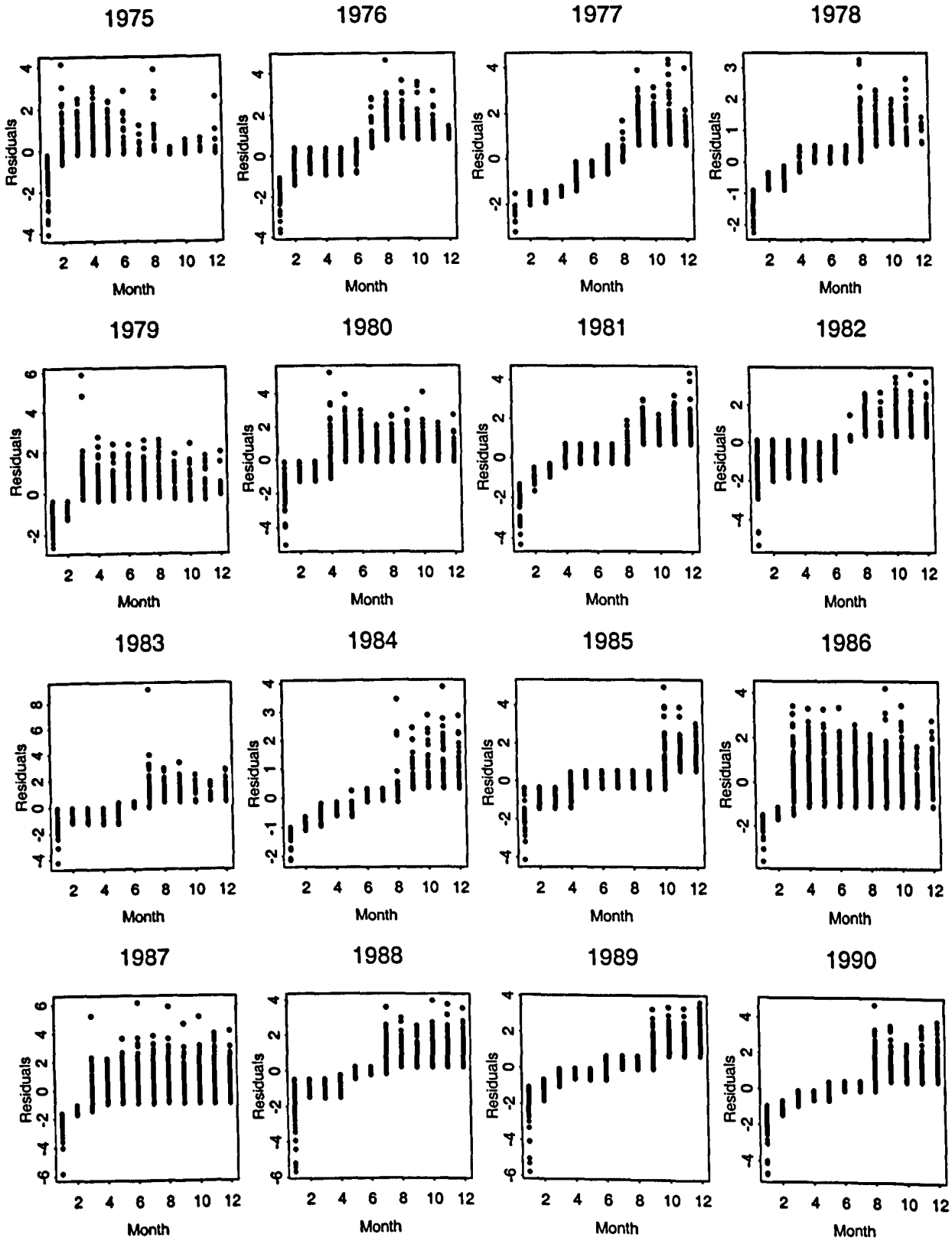


Figure 5.6: Plots of residuals against ordered values of month for the count model from Chapter 2, for 1975–1997.

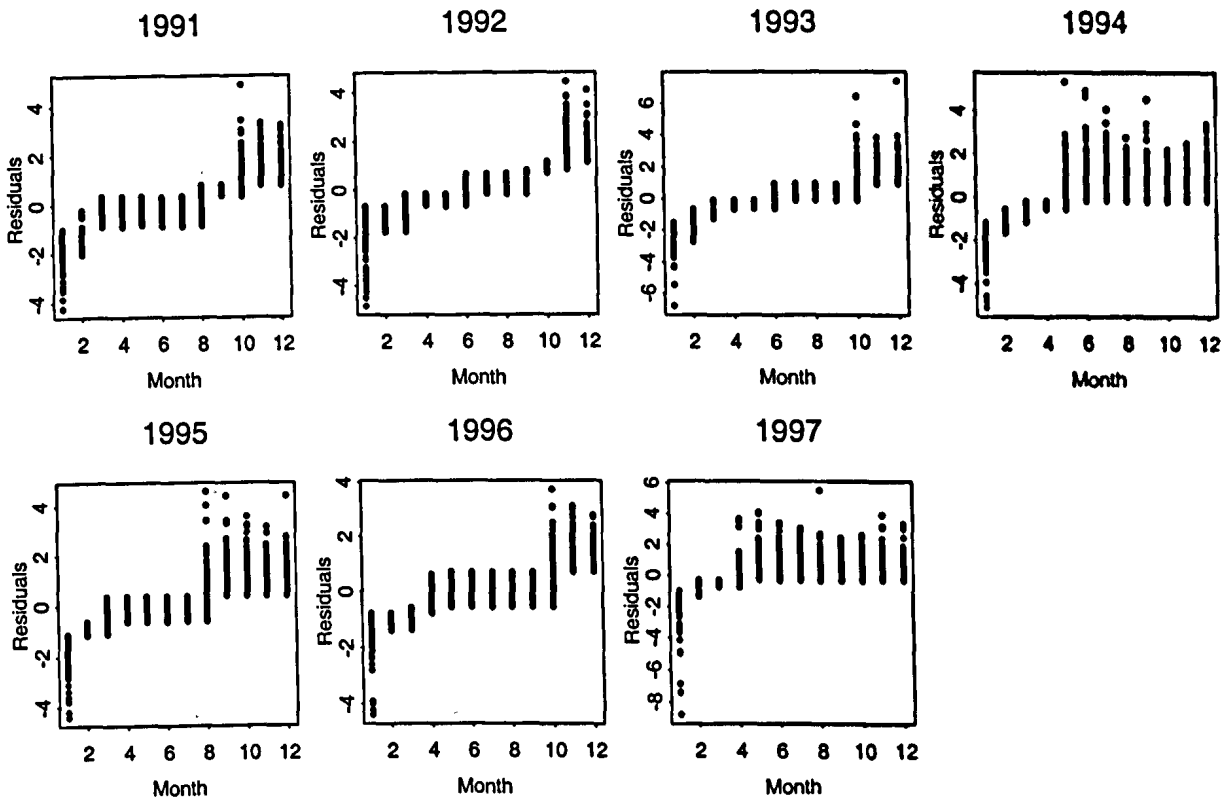


Figure 5.6: Plots of residuals against ordered values of month for the count model from Chapter 2, for 1975–1997.[continued]

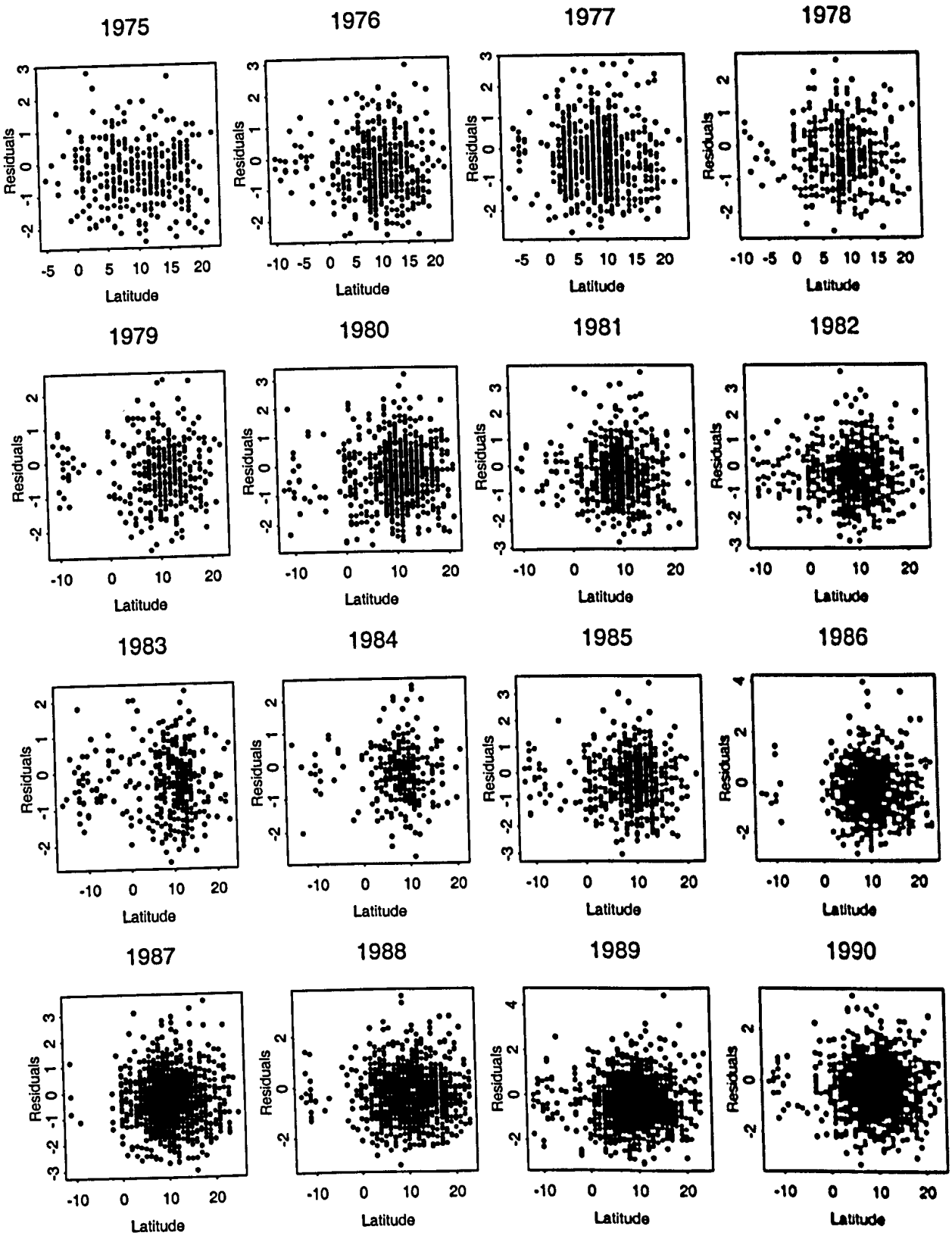


Figure 5.7: Plots of residuals against ordered values of latitude for the mean school size model from Chapter 2, for 1975–1997.

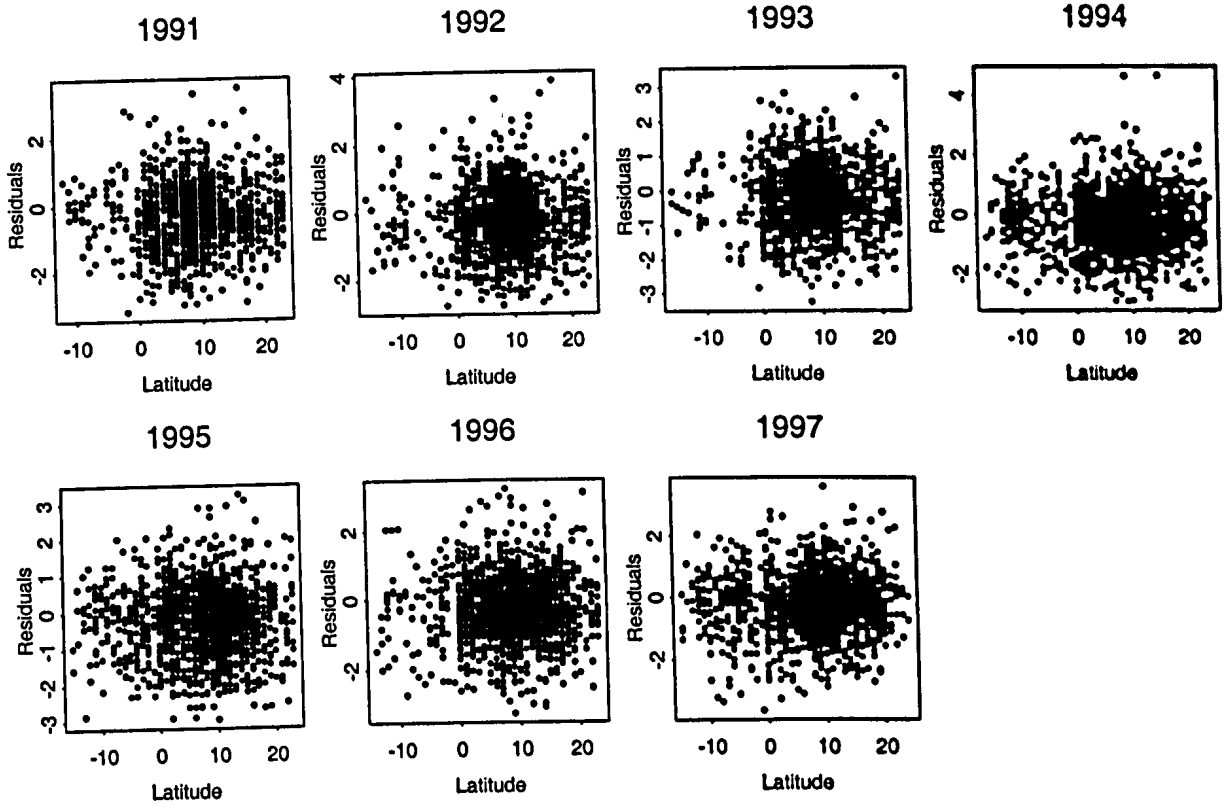


Figure 5.7: Plots of residuals against ordered values of latitude for the mean school size model from Chapter 2, for 1975–1997.[continued]

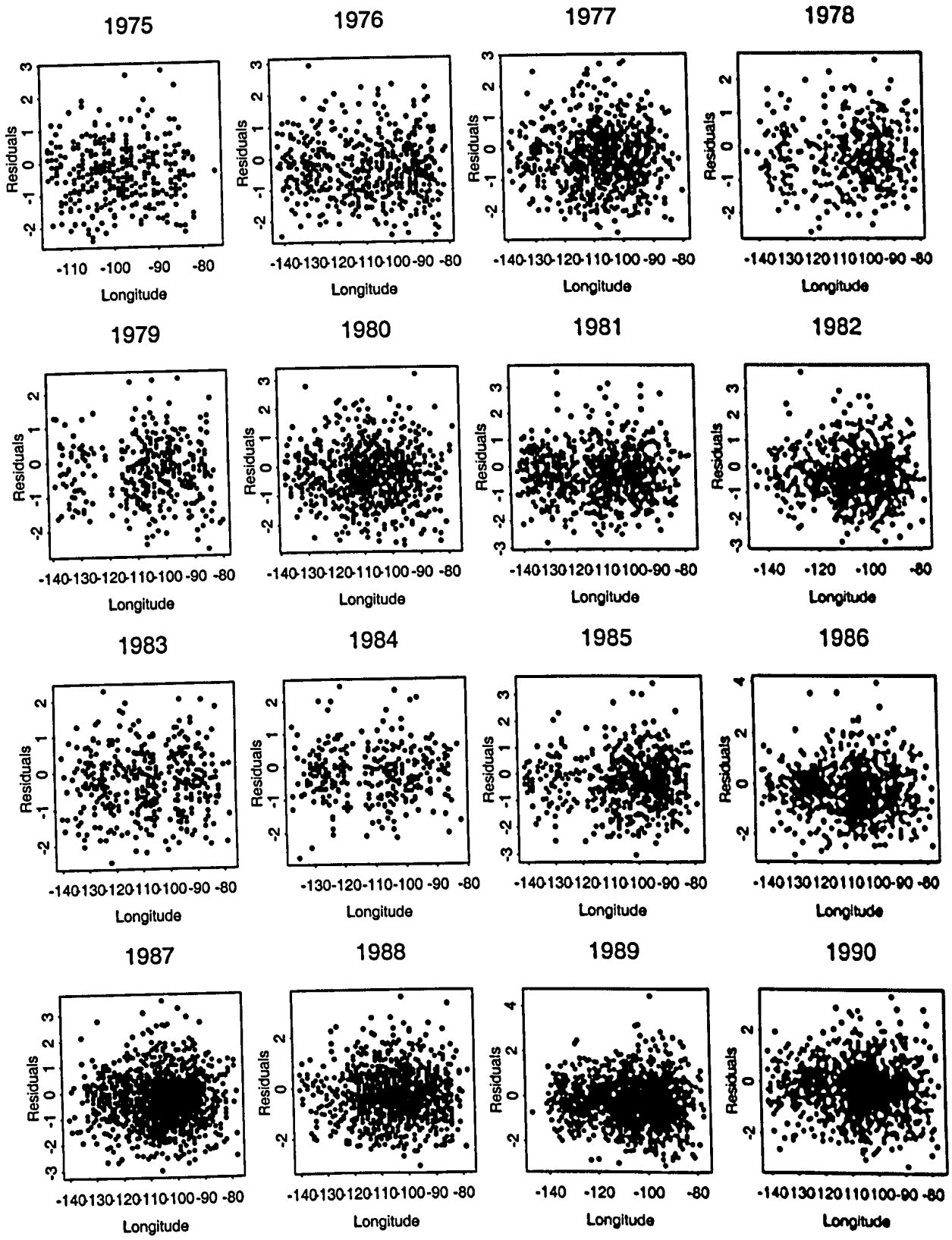


Figure 5.8: Plots of residuals against ordered values of longitude for the mean school size model from Chapter 2, for 1975–1997.

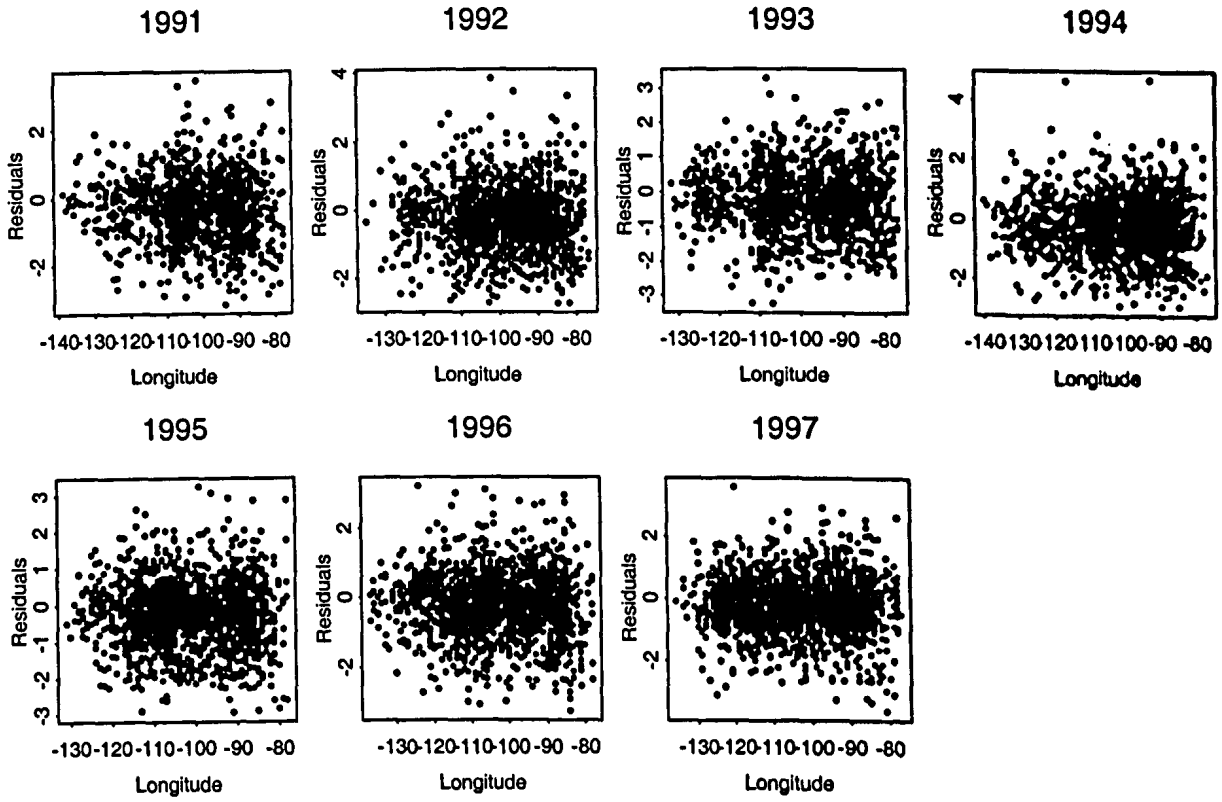


Figure 5.8: Plots of residuals against ordered values of longitude for the mean school size model from Chapter 2, for 1975–1997.[continued]

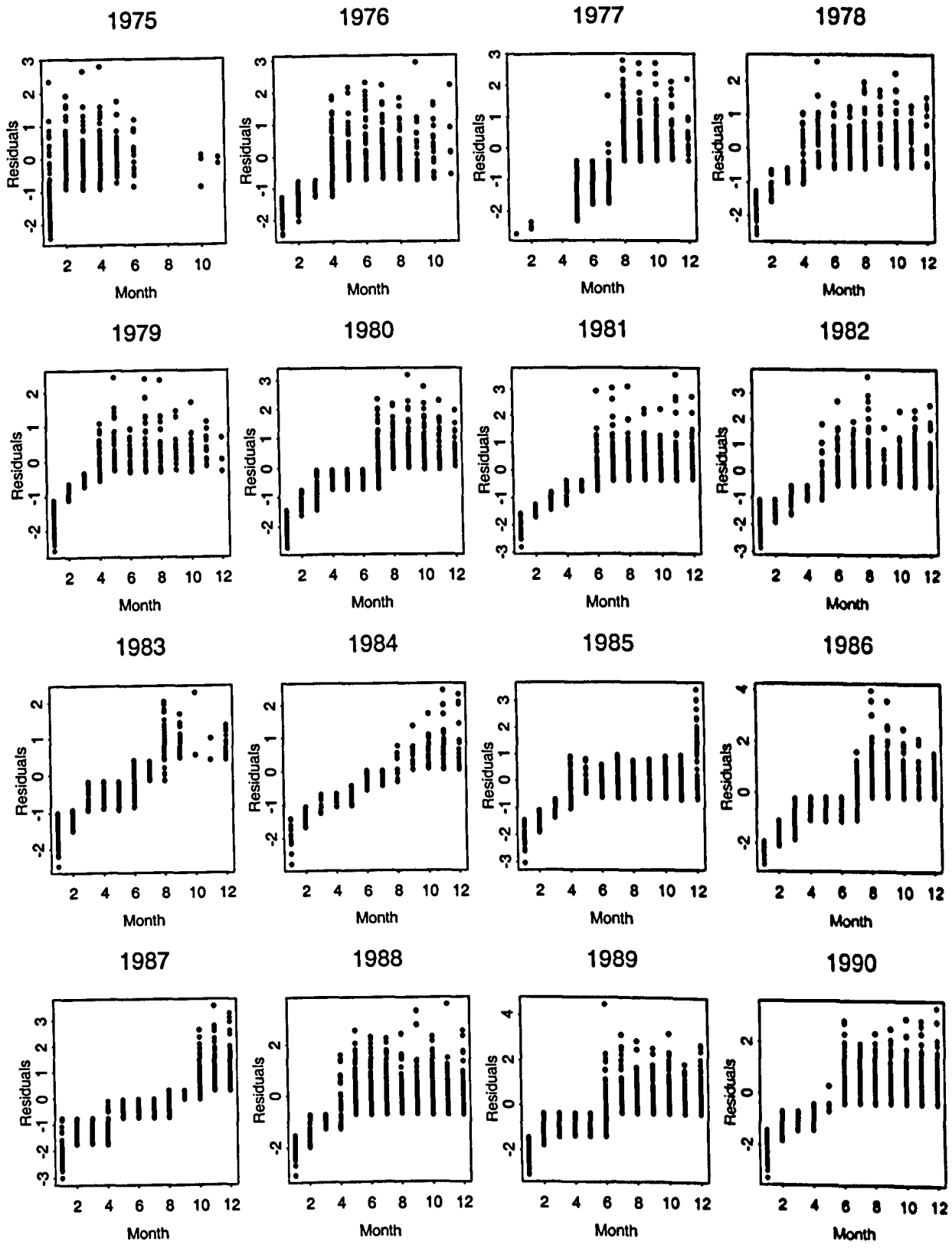


Figure 5.9: Plots of residuals against ordered values of month for the mean school size model from Chapter 2, for 1975–1997.

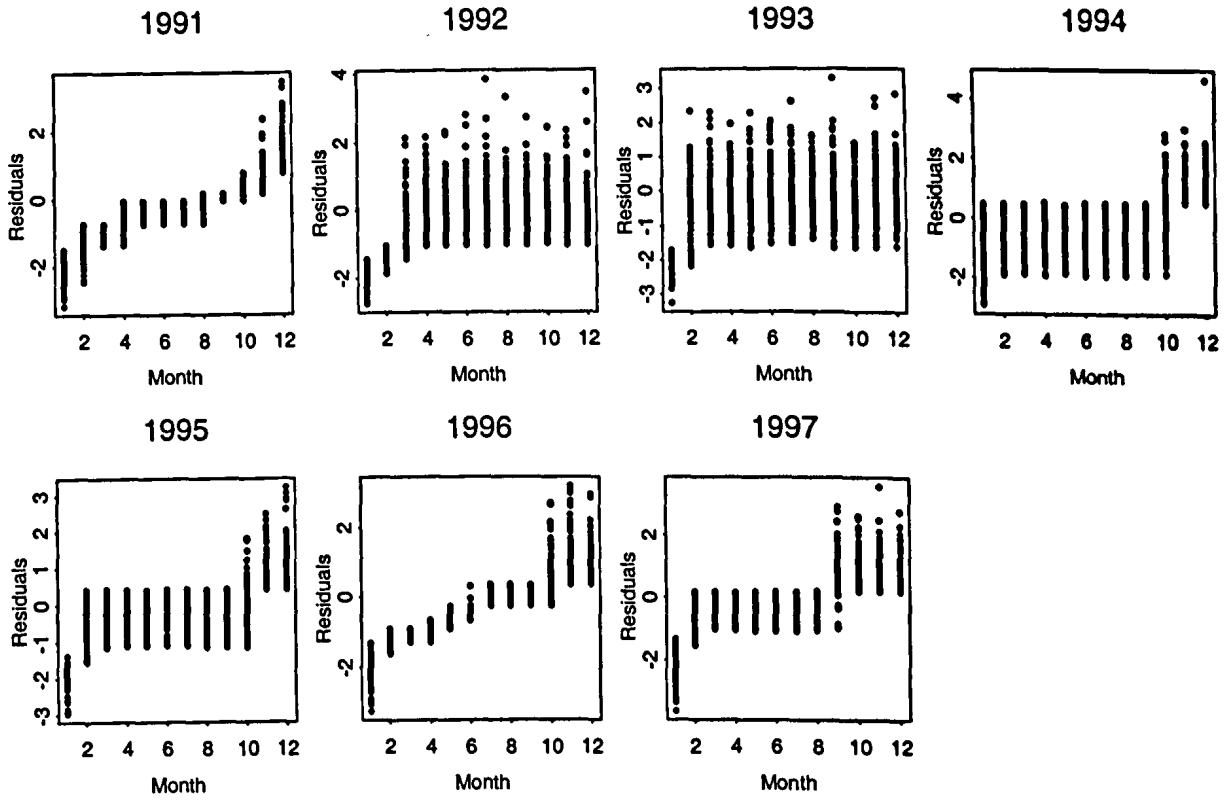


Figure 5.9: Plots of residuals against ordered values of month for the mean school size model from Chapter 2, for 1975–1997.[continued]

5.4 Trend estimation

Spatial variability in trend estimates over the small areas diminished our ability to discern any patterns in the trends. This variability may be reduced via a more integrated approach, in which both spatial and temporal variation is modelled simultaneously. This would involve the modelling of data from all years at the same time. However, this would require considerable simplification of the data in order to make the approach computationally feasible. Additional research, combined with simulations, are required to address this issue.

Another issue that remains to be addressed is the need for a more objective means of testing for differences between trend estimates from different small areas. Standard tests for regression are no longer applicable since we are not dealing with linear relationships in the trends. However, since our ability to detect significant trends in relative abundance appears to be limited, it does not seem possible to develop these objective criteria until we have improved our ability and power to detect significant trends within each small area.

Bibliography

Akaike, H. (1973) *Information theory and an extension of the maximum likelihood principle. In* Petran, B. N. and Csàaki, F. (Eds) 2nd International Symposium on Information Theory. Akademiai Kiàdo, Budapest, Hungary. pp. 267–281.

Anganuzzi, A. A. and Buckland, S. T. (1989) Reducing bias in trends in dolphin relative abundance, estimated from tuna vessel data. *Report of the International Whaling Commission* 39: 323–334.

Anganuzzi, A. A. and Buckland, S. T. (1993) Post-stratification as a bias reduction technique. *Journal of Wildlife Management* 57: 827–834.

Anganuzzi, A. A. and Buckland, S. T. (1994) Relative abundance of dolphins associated with tuna in the eastern Pacific Ocean: Analysis of 1992 data. *Report of the International Whaling Commission* 44: 361–366.

Anganuzzi, A. A., Buckland, S. T. and Cattanach, K. L. (1991) Relative abundance of dolphins associated with tuna in the eastern tropical Pacific, estimated from tuna vessel sightings data for 1988 and 1989. *Report of the International Whaling Commission* 41: 497–506.

Anganuzzi, A. A., Cattanach, K. L. and Buckland, S. T. (1992) Relative abundance of dolphins associated with tuna in the eastern tropical Pacific in 1990 and trends since 1975, estimated from tuna vessel sightings data. *Report of the International Whaling Commission* 42: 541–546.

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

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

Augustin, N. H., Borchers, D. L., Clarke, E. D., Buckland, S. T. and Walsh, M. (1998) Spatiotemporal modelling for the annual egg production method of stock assessment using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences* **55**: 2608–2621.

Barlow, J. (1984) Reproductive seasonality in pelagic dolphins (*Stenella* spp.): Implications for measuring rates. *Report of the International Whaling Commission* (Special Issue 6): 191–198.

Barlow, J. and Hohn, A. (1984) *Interpreting spotted dolphin age distributions.* U.S. Department of Commerce, NOAA Technical Memorandum NMFS 48. 22pp.

Beavers, S. C. and Ramsey, F. L. (1998) Detectability analysis in transect surveys. *Journal of Wildlife Management* **62**: 948–957.

Borchers, D. L. (1996) *Line transect estimation with uncertain detection on the trackline.* Unpublished PhD Dissertation, University of Cape Town. 233pp.

Borchers, D. L., Richardson, A. and Motos, L. (1997*a*) Modelling the spatial distribution of fish eggs using generalised additive models. *Ozeanografika* **2**: 103–120.

Borchers, D. L., Buckland, S. T., Priede, I. G. and Ahmadi, S. (1997*b*) Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences* **54**: 2727–2742.

Borchers, D. L., Zucchini, W. and Fewster, R. M. (1998*a*) Mark-recapture models for line transect surveys. *Biometrics* **54**: 1207–1220.

Borchers, D. L., Buckland, S. T., Goedhart, P. W., Clarke, E. D. and Hedley, S. L. (1998*b*) Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics* **54**: 1221–1237.

Brown, D. G. (1994) Predicting vegetation types at treeline using topography and biophysical disturbance variables. *Journal of Vegetation Science* **5**: 641–656.

Buckland, S. T. (1992*a*) Fitting density functions with polynomials. *Applied Statistics* **41**: 63–76.

Buckland, S. T. (1992*b*) Maximum likelihood fitting of Hermite and simple polynomial densities. *Applied Statistics* **41**: 241–266.

Buckland, S. T. and Anganuzzi, A. A. (1988*a*) Comparison of smearing methods in the analysis of minke sightings data from IWC/IDCR Antarctic cruises. *Report of the International Whaling Commission* **38**: 257–263.

Buckland, S. T. and Anganuzzi, A. A. (1988*b*) Estimated trends in abundance of dolphins associated with tuna in the eastern tropical Pacific. *Report of the International Whaling Commission* **38**: 411–437.

Buckland, S. T. and Turnock, B. J. (1992) A robust line transect method. *Biometrics* **48**: 901–909.

Buckland, S. T., Cattanach, K. L. and Anganuzzi, A. A. (1992) Estimating trends in abundance of dolphins associated with tuna in the eastern tropical Pacific Ocean, using sightings data collected on commercial tuna vessels. *Fishery Bulletin* **90**: 1–12.

Buckland, S. T., Anderson, D. R., Burnham, K. P. and Laake, J. L. (1993) *Distance sampling: Estimating abundance of biological populations.* Chapman & Hall, London. 446pp.

Burnham, K. P., Anderson, D. R. and Laake, J. L. (1980) Estimation of density from line transect sampling of animal populations. *Wildlife Monograph* **72**. 202pp.

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. *Report of the International Whaling Commission* **32**: 883–888.

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. *Report of the International Whaling Commission* **38**: 301–313.

Caughley, G. (1974) Bias in aerial survey. *Journal of Wildlife Management* **38**: 921–933.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. (1983) *Graphical methods for data analysis.* Wadsworth International Group, Belmont, California. 395pp.

Chapman, D. G. (1951) Some properties of the hypergeometric distribution with applications to zoological censuses. *University of California Publications in Statistics* **1**: 131–160.

Chen, S. X. (1996) Studying school size effects in line transect sampling using the kernel method. *Biometrics* **52**: 1283–1294.

Chen, S. X. (1998) Measurement errors in line transect surveys. *Biometrics* **54**: 899–908.

Chivers, S. J. and Myrick Jr., A. C. (1993) Comparison of age at sexual maturity and other reproductive parameters for two stocks of spotted dolphin, *Stenella attenuata*. *Fishery Bulletin* **91**: 611–618.

Cleveland, W. S. (1979) Robust locally-weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**: 829–836.

Cleveland, W. S. (1993) *Visualizing data.* Hobart Press, New Jersey. 360pp.

Cooke, J. G. and Leaper, R. (1998) A model to account for the rounding of recorded angles and distances in shipborne sightings surveys. Paper SC/50/RMP19 presented to the Scientific Committee of the International Whaling Commission, May 1998. 11pp. (unpublished)

Crain, B. R., Burnham, K. P., Anderson, D. R. and Laake, J. L. (1979) Nonparametric estimation of population density for line transect sampling using Fourier series. *Biometrics* **21**: 731–748.

Davison, A. C., Hinkley, D. V. and Schechtman, E. (1986) Efficient bootstrap simulation. *Biometrika* **73**: 555–566.

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

DeMaster, D. P., Edwards, E. F., Wade, P. and Sisson, J. E. (1992) *Status of dolphin stocks in the eastern tropical Pacific. In* McCullough, D. R. and Barrett, R. H. (Eds) Wildlife 2001: Populations. Elsevier Science Publishers Ltd., London. pp. 1038–1050.

DeMaster, D. P. and Sisson, J. E. (1992) *Minutes from a workshop on status of porpoise stocks in the eastern tropical Pacific, with special emphasis on the period 1985–1990.* U.S. Department of Commerce, NOAA Administrative Report NMFS 21. 69pp.

Dizon, A. E., Perrin, W. F. and Akin, P. A. (1994) *Stocks of dolphins (Stenella* spp. *and Delphinus delphis) in the eastern tropical Pacific: A Phylogeographic classification.* U.S. Department of Commerce, NOAA Technical Report NMFS 119. 20pp.

Dizon, A. E., Southern, S. O. and Perrin, W. F. (1991) Molecular analysis of mtDNA types in exploited populations of spinner dolphins (*Stenella longirostris*). *Report of the International Whaling Commission* (Special Issue 13): 183–202.

Dizon, A. E., Lockyer, C., Perrin, W. F., DeMaster, D. P. and Sisson, J. (1992) Rethinking the stock concept: A phylogeographic approach. *Conservation Biology* **6**: 24–36.

Douglas, M. E., Schnell, G. D. and Hough, D. J. (1986) Variation in spinner dolphins *Stenella longirostris* from the eastern tropical Pacific Ocean: Sexual dimorphism in cranial morphology. *Journal of Mammalogy* **67**: 537–544.

Douglas, M. E., Schnell, G. D., Hough, D. J. and Perrin, W. F. (1992) Geographic variation in cranial morphology of spinner dolphins *Stenella longirostris* in the eastern tropical Pacific Ocean. *Fishery Bulletin* **90**: 54–76.

Drummer, T. D. and McDonald, L. L. (1987) Size bias in line transect sampling. *Biometrics* **44**: 13–21.

Efron, B. and Tibshirani, R. J. (1993) *An Introduction to the bootstrap.* Chapman & Hall, London. 436pp.

Fancy, S. G. (1997) A new approach for analyzing bird densities from variable circular-plot counts. *Pacific Science* **51**: 107–114.

Fewster, R. M. and Buckland, S. T. (1996) *Quantifying population trends from CBC counts.* Report to the BTO, RUWPA, Mathematical Institute, North Haugh, St Andrews, Fife, Scotland. 28pp.

Fiedler, P. C. (1992) *Seasonal climatologies and variability of eastern tropical Pacific surface waters.* U.S. Department of Commerce, NOAA Technical Report NMFS 109. 65pp.

Fiedler, P. C. and Reilly, S. B. (1994) Interannual variability of dolphin habitats in the eastern tropical Pacific. II: Effects on abundances estimated from tuna vessel sightings, 1975–1990. *Fishery Bulletin* **92**: 451–463.

Fiedler, P. C., Chavez, F. P., Behringer, D. W. and Reilly, S. B. (1992) Physical and biological effects of Los Niños in the eastern tropical Pacific, 1986–1989. *Deep-Sea Research* **39**: 199–219.

Forney, K. A. (1997) *Patterns of variability and environmental models of relative abundance for California cetaceans.* Unpublished PhD Dissertation, University of California, San Diego. 130pp.

Forney, K. A. (1999) Trends in harbor porpoise abundance off central California, 1986-1995: Evidence for interannual changes in distribution? *Journal of Cetacean Research and Management* **1**: 73–80.

Forney, K. A. (2000) Environmental models of cetacean abundance: Reducing uncertainty in population trends. *Conservation Biology* **14**: 1271–1286.

Forney, K. A. and Barlow, J. (1993) Preliminary winter abundance estimates for cetaceans along the California coast based on a 1991 aerial survey. *Report of the International Whaling Commission* **43**: 407–415.

Gauldie, R. W. (1991) Taking stock of genetic concepts in fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences* **48**: 722–731.

Green, P. J. and Silverman, B. W. (1994) *Nonparametric regression and generalised linear models: A roughness penalty approach.* Chapman & Hall, London. 182pp.

Gunnlaugsson, T. and Sigurjónsson, J. (1990) NASS-87: Estimation of whale abundance based on observations made onboard Icelandic and Faroese survey vessels. *Report of the International Whaling Commission* **40**: 571–580.

Hall, M. A. (1998) An ecological view of the tuna-dolphin problem: Impact and trade-offs. *Reviews in Fish Biology and Fisheries* **8**: 1–34.

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. F. and Øien, N. (1995) *Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters.* SCANS Final Report Life 92-2/UK/027, October 1995. 240pp.

Hansen, M. H., Madow, W. G. and Tepping, B. J. (1983) An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association* **78**: 776–793.

Hastie, T. J. and Tibshirani, R. J. (1990) *Generalized additive models.* Chapman & Hall, London. 335pp.

Hayes, R. J. and Buckland, S. T. (1983) Radial-distance models for the line-transect method. *Biometrics* **39**: 29–42.

Hedley, S. L. (2000) *Modelling heterogeneity in cetacean surveys.* Unpublished PhD Dissertation, University of St Andrews. 139pp.

Hedley, S. L., Buckland, S. T. and Borchers, D. L. (1999) Spatial modelling from line transect data. *Journal of Cetacean Research and Management* **1**: 255–264.

Hiby, A. R. (1986) Results of a hazard rate model relevant to experiments on the 1984/1985 IDCR minke whale assessment cruise. *Report of the International Whaling Commission* **36**: 497–498.

Hohn, A. A. and Hammond, P. S. (1985) Early postnatal growth of the spotted dolphin, *Stenella attenuata*, in the offshore eastern tropical Pacific. *Fishery Bulletin* **83**: 553–566.

Hohn, A. A., Chivers, S. J. and Barlow, J. (1985) Reproductive maturity and seasonality of male spotted dolphins, *Stenella attenuata*, in the eastern tropical Pacific. *Marine Mammal Science* **1**: 273–293.

Horvitz, D. G. and Thompson, D. J. (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**: 663–685.

IATTC (1994) *Annual Report of the Inter-American Tropical Tuna Commission, 1992.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IATTC (1995) *Annual Report of the Inter-American Tropical Tuna Commission, 1993.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IATTC (1996) *Annual Report of the Inter-American Tropical Tuna Commission, 1994.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IATTC (1997) *Annual Report of the Inter-American Tropical Tuna Commission, 1995.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IATTC (1998) *Annual Report of the Inter-American Tropical Tuna Commission, 1996.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IATTC (1999) *Annual Report of the Inter-American Tropical Tuna Commission, 1997.* Inter-American Tropical Tuna Commission, La Jolla, California, USA.

IWC (1992) Annex I: Background to the development of revised management procedures. *Report of the International Whaling Commission* **42**: 236–243.

IWC (1993*a*) Annex E: Report of the Sub-Committee on southern hemisphere baleen whales. *Report of the International Whaling Commission* **43**: 104–114.

IWC (1993*b*) Annex H: Draft specification for the calculation of catch limits in a revised management procedure (RMP) for baleen whales. *Report of the International Whaling Commission* **43**: 146–152.

Joseph, J. (1994) The tuna-dolphin controversy in the eastern Pacific Ocean: Biological, economic, and political impacts. *Ocean Development and International Law* **25**: 1–30.

Kullback, S. and Leibler, R. A. (1951) On information and sufficiency. *Annals of Mathematical Statistics* **22**: 79–86.

Laake, J. L., Buckland, S. T., Anderson, D. R. and Burnham, K. P. (1993) *DISTANCE User's Guide.* Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO 80523, USA.

Lennert-Cody, C., Buckland, S. T. and Marques, F. F. C. (In prep) Trends in dolphin abundance estimated from fisheries data: A cautionary note.

Mack, Y. P. and Quang, P. X. (1998) Kernel methods in line and point transect sampling. *Biometrics* **54**: 606–619.

McCullagh, P. and Nelder, J. A. (1989) *Generalized linear models.* Chapman & Hall, London. 511pp.

Moody, J. E. (1991) *Note on generalization, regularization, and architecture selection in nonlinear learning systems. In* Juang, B. H., Kung, S. Y. and Kamm, C. A. (Eds) Neural Networks for Signal Processing. IEEE Press, Piscataway, NJ. pp. 1–10.

Moody, J. E. (1992) *The effective number of parameters: An analysis of generalization and regularization in nonlinear learning systems. In* Moody, J. E., Hanson, S. J. and Lippmann, R. P. (Eds) Advances in Neural Information Processing Systems 4. Morgan Kaufmann Publishers, San Mateo, CA. pp. 847–854.

Murata, N., Yoshizawa, S. and Amari, S. (1994) Network Information Criterion – Determining the number of hidden units for an artificial neural network model. *IEEE Transactions on Neural Networks* **5**: 865–872.

Otto, M. C. and Pollock, K. H. (1990) Size bias in line transect sampling: A field test. *Biometrics* **46**: 239–245.

Palka, D. L. (1993) *Estimating density of animals when assumptions of line transect survey are violated.* Unpublished PhD Dissertation, University of California, San Diego. 169pp.

Perrin, W. F. (1984) Patterns of geographical variation in small cetaceans. *Acta Zoologica Fennica* **172**: 137–140.

Perrin, W. F. (1990) Subspecies of *Stenella longirostris* (Mammalia: Cetacea: Delphinidae). *Proceedings of the Biological Society of Washington* **103**: 453–463.

Perrin, W. F. and Henderson, J. R. (1984) Growth and reproductive rates in two populations of spinner dolphins, *Stenella longirostris*, with different histories of exploitation. *Report of the International Whaling Commission* (Special Issue 6): 417–430.

Perrin, W. F. and Reilly, S. B. (1984) Reproductive parameters of dolphins and small whales of the Family Delphinidae. *Report of the International Whaling Commission* (Special Issue 6): 97–133.

Perrin, W. F., Akin, P. A. and Kashiwada, J. V. (1991) Geographic variation in external morphology of the spinner dolphin *Stenella longirostris* in the eastern Pacific and implications for conservation. *Fishery Bulletin* **89**: 411–428.

Perrin, W. F., Evans, W. E. and Holts, D. B. (1979) *Movements of pelagic dolphins (Stenella* spp.*) in the eastern tropical Pacific as indicated by results of tagging, with summary of tagging operations, 1969–76.* U.S. Department of Commerce, NOAA Technical Report NMFS 737. 14pp.

Perrin, W. F., Schnell, G. D., Hough, D. J., Gilpatrick Jr., J. W. and Kashiwada, J. V. (1994) Reexamination of geographic variation in cranial morphology of the pantropical spotted dolphin, *Stenella attenuata*, in the eastern Pacific. *Fishery Bulletin* **92**: 324–346.

Quang, P. X. (1991) A nonparametric approach to size-biased line transect sampling. *Biometrics* **47**: 269–279.

Ramsey, F. L., Wildman, V. W. and Engbring, J. (1987) Covariate adjustments to effective area in variable-area wildlife surveys. *Biometrics* **43**: 1–11.

Reilly, S. B. (1990) Seasonal changes in distribution and habitat differences among dolphins in the eastern tropical Pacific. *Marine Ecology Progress Series* **66**: 1–11.

Reilly, S. B. and Fiedler, P. C. (1994) Interannual variability of dolphin habitats in the eastern tropical Pacific. I: Research vessel surveys, 1986–1990. *Fishery Bulletin* **92**: 434–450.

Ripley, B. D. (1987) *Stochastic simulation.* John Wiley, New York. 237pp.

Ripley, B. D. (1996) *Pattern recognition and neural nertworks.* Cambridge University Press, Cambridge. 403pp.

Schnell, G. D., Douglas, M. E. and Hough, D. J. (1986) Geographic patterns of variation in offshore spotted dolphins (*Stenella attenuata*) of the eastern tropical Pacific Ocean. *Marine Mammal Science* **2**: 186–213.

Schweder, T. (1990) Independent observer experiments to estimate the detection function in line transect surveys of whales. *Report of the International Whaling Commission* **40**: 349–355.

Seber, G. A. F. (1977) *Linear regression analysis.* John Wiley & Sons, London. 465pp.

Seber, G. A. F. (1982) *The estimation of animal abundance.* Griffin, London. 654pp.

Smith, T. D. (1983) Changes in size of three dolphin (*Stenella* spp.) populations in the eastern tropical Pacific. *Fishery Bulletin* **81**: 1–13.

Swartzman, G. (1997) Analysis of the summer distribution of fish schools in the Pacific Eastern Boundary Current. *ICES Journal of Marine Science* **54**: 105–116.

Swartzman, G., Huang, C. and Kaluzny, S. (1992) Spatial analysis of Bering Sea groundfish survey data using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences* **49**: 1366–1378.

Swartzman, G., Silverman, E. and Williamson, N. (1995) Relating trends in walleye pollock (*Theragra chalcogramma*) abundance in the Bering Sea to environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences* **54**: 105–116.

Swartzman, G., Stuetzle, W., Kulman, K. and Powojowski, M. (1994) Relating the distribution of pollock schools in the Bering Sea to environmental factors. *ICES Journal of Marine Science* **51**: 481–492.

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

Venables, W. N. and Ripley, B. D. (1994) *Modern applied statistics with S-Plus.* Springer-Verlag, London. 462pp.

Wade, P. R. (1993) Estimation of historical population size of the eastern spinner dolphin *Stenella longirostris orientalis. Fishery Bulletin* **91**: 775–787.

Wade, P. R. (1994) *Abundance and population dynamics of two eastern Pacific dolphins, Stenella attenuata and Stenella longirostris orientalis.* Unpublished PhD Dissertation, University of California, San Diego. 255pp.

Wade, P. R. and Angliss, R. P. (1997) *Guidelines for assessing marine mammal stocks.* Report of the GAMMS Workshop, April 3–5, 1996, Seattle, Washington, USA. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-OPR-12.

Wade, P. R. and Gerrodette, T. (1992) Estimates of dolphin abundance in the eastern tropical Pacific: Preliminary analysis of five years of data. *Report of the International Whaling Commission* **42**: 533–539.

Welch, D. W., Chigirinsky, A. I. and Ishida, Y. (1995) Upper thermal limits on the oceanic distribution of Pacific salmon (*Oncorhynchus* spp.) in the spring. *Canadian Journal of Fisheries and Aquatic Sciences* **52**: 489–503.

Wood, S. N. (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society B* **62**: 413–428.

Wyrtki, K. (1966) Oceanography of the eastern equatorial Pacific Ocean. *Oceanography and Marine Biology Annual Review* **4**: 33–68.

Appendix A

Introduction

In the eastern tropical Pacific (ETP) Ocean, yellowfin tuna (*Thunnus albacares*) commonly concentrate under schools of spotted (*Stenella attenuata*), spinner (*S. longirostris* spp.) and, to a lesser extent, common (*Delphinus delphis* spp.) dolphins. This relationship has long been exploited by tuna purse seine fishermen, who often use dolphins as a cue to locate the tuna and encircle dolphin schools to retrieve the fish underneath them (for a review of the history of the fishery see Joseph (1994) and Hall (1998)). Incidental mortality of dolphins during purse seine operations has negatively affected populations of spotted and spinner dolphins (Smith 1983). Incidental takes of striped (*S. coeruleoalba*) and common dolphins also occur, but the evaluation of the impact of this mortality on their abundance has been problematic (Buckland and Anganuzzi 1988*b*; Buckland *et al.* 1992).

Since 1974 the National Marine Fisheries Service of the United States has placed observers on board US-registered tuna vessels, to monitor dolphin mortality resulting from interaction with the tuna fishery. Coverage of the international fleet (including US-registered vessels) has been carried out by the Inter-American Tropical Tuna Commission since 1979. Dolphin sightings data gathered by tuna vessel observers are used to evaluate the status of dolphin populations. Annual estimates of dolphin relative abundance obtained using line transect methods are used to assess trends. A description of the various methodologies can be found in Buckland and Anganuzzi (1988*b*), Anganuzzi and Buckland (1989), and Buckland *et al.* (1992).

The tuna vessel observer data has been used in all examples presented in this thesis. Here a brief description of data collection procedures relevant to the line transect methodology is presented; for additional information see Joseph (1994), Hall (1998), and the references cited in those papers.

Search effort

The tuna vessels from which the observers collect dolphin sightings data are part of a commercial enterprise, and hence can be seen as 'platforms of opportunity'. The vessels do not follow any survey design; instead the vast majority tends to search over areas of known high dolphin densities, where the larger dolphin schools (which are normally associated with the larger schools of tuna) are expected to occur. As schools of yellowfin tuna are commonly associated with schools of offshore spotted dolphins, or with mixed schools of offshore spotted and spinner dolphins, the vessels concentrate in areas where those species are abundant. A smaller proportion of the vessels search for free-swimming schools of tuna, or for schools associated with floating objects (Hall 1998), and hence cover the remaining regions of the ETP where the density of offshore spotted and spinner dolphins is not as high. Due to this non-random distribution of search effort, it is not possible to obtain estimates of absolute dolphin abundance using the tuna vessel observer data, as such estimates would be biased upwards. Instead, the data are used to estimate dolphin *relative* abundance, from which trends in the populations can be ascertained provided the biases do not change over time.

Search effort is carried out throughout most of the time the vessels are at sea, weather permitting, during the day, and whether or not the observer is also on effort. Only when the vessels begin to set on a school of tuna, regardless of whether it is associated with dolphins or not, that search effort ceases.

Since the implementation of the observer program in the mid-1970s, a number of technological developments has taken place in the tuna fishery, with direct implications for the way the data are collected (Lennert-Cody *et al.* In prep). Initially most of the search effort was carried out by 20x binoculars, but in the late 1980s the use of high resolution radar was implemented. Because the radar is able to detected seabirds flying above dolphin schools, it allowed sightings to be made at greater distances. At about the same time there was an increase in the use of helicopters to search for both schools of dolphin and schools of tuna, and now most, if not all, tuna vessels have a helicopter on board.

Political pressures have also affected the way the tuna fishery operates. The adoption of the 'dolphin-safe' policy, in which vessels were labelled 'dolphin-safe' if they did not intentionally encircle any dolphin schools; the ban by the US government on the importation of tuna caught by sets made on dolphins; and the imposition of limits on dolphin mortality associated with the fishery; all these affected how the vessels searched for dolphin schools

(see Lennert-Cody *et al.* (In prep) for a review).

Data collection

Each vessel has a unique cruise number associated with its current trip, and that number is associated with all the data recorded for that particular vessel during that trip. Whenever the vessel is on effort, *i.e.* searching for dolphins or tuna, it is noted whether or not the observer is also on effort. Search for dolphins is primarily carried out by the crew of the vessels, though the observers also make sightings.

On each vessel the observer records the position of the vessel whenever it changes course. Sea condition (Beaufort) is also recorded whenever it changes, along with the time and position of the vessel. The date is always recorded together with all the data.

When a dolphin sighting is made, the time and position of the vessel are noted, together with the initial sighting angle and radial distance. It is not always possible to record the exact position of all sightings, so an accuracy code, indicating the accuracy of the recorded position (e.g. position known, or known to the nearest minute, etc.) is also entered. For each sighting the sighting method and sighting cue are recorded (see Table below). An estimate of the size of the school and its species composition are also noted, both by the observer and the crew separately. If the sighting leads to a set, this is noted, and at the time of the set both the observer and the crew again give separate estimates of the school size and species composition.

Sighting method	Sighting cue
By 20x from crow's nest	Birds
From crow's nest	Splashes
By 20x binoculars	Log, breezer, or other cue
From helicopter	The dolphins themselves
Zero people on 20x binoculars	
One people on 20x binoculars	
Two people on 20x binoculars	
Three people on 20x binoculars	
Four people on 20x binoculars	
Five people on 20x binoculars	
By radar	

Data handling

To minimise any biases resulting from the data collection procedures, the data are filtered in a number of ways, described in detail by Buckland and Anganuzzi (1988*b*) and Anganuzzi and Buckland (1989). Here only a brief summary is presented; for more information see the papers cited.

For analyses purposes, only the search effort carried out while the observer was also on effort is included. In addition, because under Beaufort conditions greater than three the detectability of dolphin schools is affected (Buckland and Anganuzzi 1988*b*), only effort carried out in Beaufort up to three is considered.

In the early years, the vessels often turned towards the dolphin school before the observer recorded the initial sighting angle and distance. As a result, a large proportion of sightings had an angle of zero. This led to a disproportionate number of zero perpendicular distances, which caused a positive bias in the estimates of $f(0)$ obtained from these data. To minimise this bias, cruises with an average sighting angle smaller than 20° are discarded.

Although the observers occasionally make sightings, their search pattern is substantially different from that of the crew. In addition, the observers are often busy recording the data, so it is difficult to quantify their search effort. Therefore sightings made by the observers are not included in the analyses.

Estimates of school size and species composition are based on the crew's estimates obtained during a set. If that information is not available, a correction factor for the observer's estimates (in comparison with the crew's estimates) is obtained and applied to the observer's estimates. For sightings which did not lead to a set, the crew's estimates are then used; if those are not available, a correction factor is then applied to the observer's estimates.

Interpolation based on the vessel's positions is used to estimate sighting positions for which the exact latitude and longitude of the sighting are not known.