

# 8

---

## *Illustrative examples*

### **8.1 Introduction**

Several analyses of real data are presented here to illustrate line and point transect analysis, and use of program DISTANCE. Units of measurement of the original study are adhered to, to avoid quoting cutpoints as 'every 0.305 m' instead of 'every foot'. Five of the examples are line transect surveys and three are point transect studies. The first two examples are case studies in which the true density is known. The first of these is a set of line transects to estimate the number of bricks that had been placed on the bed of a region of Lake Huron, as part of a programme to assess the viability of using an underwater video system to monitor numbers of dead lake trout. This is followed by a reanalysis of the wooden stake surveys carried out as part of a student project in sagebrush-grass near Logan, Utah (Laake 1978). The third example is a comprehensive analysis of the data from annual surveys of duck nest density that have been carried out for 27 years at the Monte Vista National Wildlife Refuge in Colorado. An analysis of fin whale data from ship sightings surveys in the North Atlantic illustrates use of stratification, and the final line transect example is one for which sightings data on dolphins are collected by observers placed on board tuna fishing vessels. Thus, there is no control over the cruise track, and methods must be used that reduce the effects of the potentially strong biases in the data.

The first point transect example is an illustration of model selection and stratification on house wren data collected during surveys of riparian vegetation in South Platte River bottomland in Colorado. The second example is an analysis of data on the six species most commonly recorded in songbird surveys of the Arapaho National Wildlife Refuge, also in Colorado. In the final example, an approach for assessing effects of habitat on density using point transect sampling is described, and an analysis of binomial point transect data is carried out in which bird

## LAKE HURON BRICK DATA

density and detectability in restocked forest plantations in Wales are related to habitat succession.

### 8.2 Lake Huron brick data

Bergstedt and Anderson (1990) presented the results of 13 independent, line transect surveys of known populations of bricks placed approximately at random on the floor of portions of Lake Huron. The bricks were to simulate dead lake trout, and the purpose of the survey was to explore the feasibility of large-scale surveys of dead lake trout on the lake bottom. The distance data were taken remotely via a video camera mounted on an underwater sled and pulled along the lake bottom by a surface vessel. The data were taken in five distance categories (0, 0.333, 0.667, 1, 1.333, 1.667 m), and these cutpoints were marked on a cathode ray tube on board the vessel. The relationship between the cutpoints on the screen and the perpendicular distance on the lake bottom was defined by calibration studies before the first survey. The true density in all surveys was 89.8 bricks/ha, and this allows an evaluation of the utility of line transect sampling theory. Full details of this survey are given by Bergstedt and Anderson (1990). Our purpose here is to reanalyse the data from the 13 surveys using the methods presented in earlier chapters.

**Table 8.1** Summary of density estimates and coefficients of variation for the Lake Huron brick data under five models of the detection function (data from Bergstedt and Anderson 1990). True density = 89.8/ha. The final means are weighted by line length

Survey	HN + cos		Uni + cos		Uni + poly		HN + Herm		Haz + poly	
1	138.5	19.5	144.4	15.7	137.1	19.2	126.8	16.6	104.0	16.2
2	96.4	18.7	92.6	32.5	100.0	15.6	93.1	27.1	88.2	15.6
3	88.8	20.4	91.4	19.2	86.2	17.0	88.8	20.4	70.0	17.7
4	84.0	19.7	85.5	15.9	83.2	19.6	77.1	16.0	63.0	16.2
5	78.3	22.0	81.6	20.0	72.9	18.4	78.3	22.0	61.3	18.9
6	84.5	33.8	93.6	14.6	90.4	17.7	84.4	27.5	75.9	16.4
7	80.7	38.2	76.9	26.3	92.4	19.3	85.4	16.6	71.1	16.0
8	122.4	19.1	107.0	15.7	104.1	19.2	122.4	19.1	86.2	16.7
9	96.2	19.5	99.6	15.7	95.7	19.2	96.2	19.5	73.6	16.0
10	72.8	20.2	70.2	26.7	67.7	16.9	67.6	30.0	61.5	16.0
11	89.4	20.9	93.4	19.2	86.1	17.3	89.4	20.9	70.6	18.0
12	92.7	18.7	83.9	15.3	81.6	18.0	81.9	29.2	76.5	18.0
13	56.8	20.7	58.2	20.0	55.1	17.8	56.8	20.7	55.5	27.2
Wt. Ave.	88.7	6.2	88.3	6.0	86.5	6.0	86.2	6.8	72.4	4.7

## ILLUSTRATIVE EXAMPLES

Table 8.1 provides a summary of the estimates of density for each of the 13 surveys under each of five models: half-normal + cosine, uniform + cosine, uniform + simple polynomial, half-normal + Hermite polynomial, and the hazard-rate model + simple polynomial. Information related to model selection for the pooled data is presented in Table 8.2. AIC selects the hazard-rate model with no adjustment terms, and this model fits the grouped distance data well ( $p = 0.70$ ). The remaining four models (half-normal + cosine, uniform + cosine, uniform + polynomial and half-normal + Hermite polynomial) have similar AIC values, but fit the data less well ( $0.09 \leq p \leq 0.29$ ).

**Table 8.2** Model selection statistics for the pooled Lake Huron brick data

Key function	Adjustment	Total parameters	AIC	$p$ -value*
Half-normal	Cosine	4	1671.2	0.09
Uniform	Cosine	3	1670.3	0.29
Uniform	Polynomial	3	1670.3	0.29
Half-normal	Hermite	3	1672.7	0.06
Hazard-rate	Polynomial	2	1667.9	0.70

\* Goodness of fit test.

Surprisingly, the hazard-rate model provides the poorest estimates of mean density based on the weighted mean of the estimates (Table 8.1). The hazard-rate model fitted a very flat shoulder to the pooled data and produced estimates that were low in 12 of the 13 surveys. Performance of the other four estimators is quite good (all were slightly low due, in part, to the results of survey 13) (Table 8.1). Estimated confidence intervals for mean density covered the true value for all but the hazard-rate model.

The weighted mean of the estimates was very close to the true parameter, except for the hazard-rate model (88.7, 88.3, 86.5, 86.2 and 72.4, respectively). Any of the first four models performs well, especially when one considers that the average sample size per survey was only 45. The hazard-rate key + polynomial adjustment does more poorly in this example, and is the only model whose 95% confidence interval fails to cover the true density. Troublesome is the fact that the hazard-rate model was selected as the best of the five models by AIC in nine of the 13 data sets (surveys).

The reader is encouraged to compare these results with the original paper by Bergstedt and Anderson (1990), which includes discussion of various points relating to possible measurement errors, missing bricks in the first distance category, and potential problems with survey 13.

## WOODEN STAKE DATA

Bergstedt and Anderson (1990) note that some bricks were missed near the centreline, and that the cutpoints drawn on the cathode ray tube, although accurate in the initial calibration, were perhaps compromised by the uneven lake bottom.

### 8.3 Wooden stake data

Laake (1978) set out 150 unpainted wooden stakes ( $2.5 \times 5 \times 46$  cm) within a rectangular area of sagebrush–grass near Logan, Utah, in the spring of 1978 to examine the performance of the line transect method.

**Table 8.3** Summary of stake data taken in 1978 in a sagebrush–grass field near Logan, Utah (Laake 1978). Density in each of the 11 surveys was 37.5 stakes/ha. Cosine adjustments were added as required in modelling  $f(x)$ . For each survey,  $L = 1000$  m and  $w = 20$  m

Survey no.	Key function	Sample size	Density estimate	cv(%)	Log-based 95% confidence interval
1	Half-normal	72	37.11	19.3	(25.51, 53.99)
	Uniform		30.00	14.5	(22.63, 39.77)
2	Half-normal	48	35.18	19.9	(23.90, 51.78)
	Uniform		36.01	20.1	(24.36, 53.23)
3	Half-normal	74	28.76	15.8	(21.16, 39.10)
	Uniform		29.26	14.8	(21.92, 39.07)
4	Half-normal	59	38.31	19.1	(26.42, 55.52)
	Uniform		33.30	17.5	(23.68, 46.81)
5	Half-normal	59	34.41	19.9	(23.37, 50.66)
	Uniform		29.58	19.0	(20.47, 42.76)
6	Half-normal	72	26.38	16.2	(19.24, 36.17)
	Uniform		27.08	15.6	(19.98, 36.69)
7	Half-normal	55	34.48	19.9	(23.44, 50.72)
	Uniform		34.69	21.1	(23.06, 52.18)
8	Half-normal	61	33.31	20.2	(22.51, 49.30)
	Uniform		34.48	21.3	(22.82, 52.09)
9	Half-normal	46	28.32	21.9	(18.51, 43.31)
	Uniform		23.52	21.2	(15.60, 35.48)
10	Half-normal	43	34.16	20.1	(23.15, 50.42)
	Uniform		32.69	21.1	(21.71, 49.22)
11	Half-normal	53	29.80	17.4	(21.25, 41.80)
	Uniform		31.45	17.8	(22.23, 44.50)
Mean	Half-normal	642	32.75	3.6	(30.54, 35.11)
	Uniform		31.10	3.6	(28.99, 33.36)
Pooled	Half-normal	642	34.37	7.2	(29.86, 39.55)
	Uniform		34.60	7.5	(29.86, 40.08)

## ILLUSTRATIVE EXAMPLES

The stakes were placed in a restricted random spatial pattern such that the number of stakes was distributed uniformly as a function of distance from the line. In fact, each 2 m distance category had 15 stakes present. Stakes were driven in the ground until about 37 cm remained above ground. One stake was placed about every 7 m of transect and alternated between left and right sides of the line. Exact placement was generated randomly within the 7 m section. True density was 37.5 stakes per hectare.

A single, well-marked line ( $L = 1000$  m,  $w = 20$  m) was traversed by 11 different, independent observers. The observers were carefully instructed and supervised and fatigue was probably a minor factor as each survey could be completed by the observer in approximately 2 h. Observers traversed the line at different times, thus the data for each of the 11 surveys are independent. The number of stakes detected ( $n$ ) varied from 43 to 74, corresponding to  $\hat{P}_a$  ranging from 0.29 to 0.49 (Table 8.3). Histograms and estimated detection functions ( $\hat{g}(x)$ ) differed greatly among observers. In field studies, the detection function would also be affected by habitat type and species being surveyed. These factors affect  $n$  and make it, alone, unreliable as an index of density.

Two strategies were used for the analysis of these data. First, the data were pooled over the 11 surveys ( $n = 642$ ) and AIC was computed for five models,

Model	AIC
Half-normal + cosine	2412.9
Uniform + cosine	2415.2
Uniform + polynomial	2417.4
Half-normal + Hermite	2450.8
Hazard-rate + cosine	2416.4

Thus, the half-normal + cosine is selected as the best model. In fact, all models seem fairly satisfactory for data analysis except the half-normal + Hermite polynomial model where the LR test indicated that the first adjustment term (for kurtosis) was not required ( $p = 0.617$ ). There are options in DISTANCE, such as **SELECT = forward** (which yields AIC = 2415.1 for this model) or **LOOKAHEAD = 2** (also giving AIC = 2415.1), that allow the user to avoid such poor model fits (Section 8.7). Estimates of density are shown in Table 8.3 under the two best models, as suggested by the AIC criterion. The estimates of density are quite similar between these two models.

While  $n$  varied widely, estimates of density varied from only 26.38 to 38.31, using two models for the detection function (half-normal and uniform key functions with cosine adjustment terms). Confidence inter-

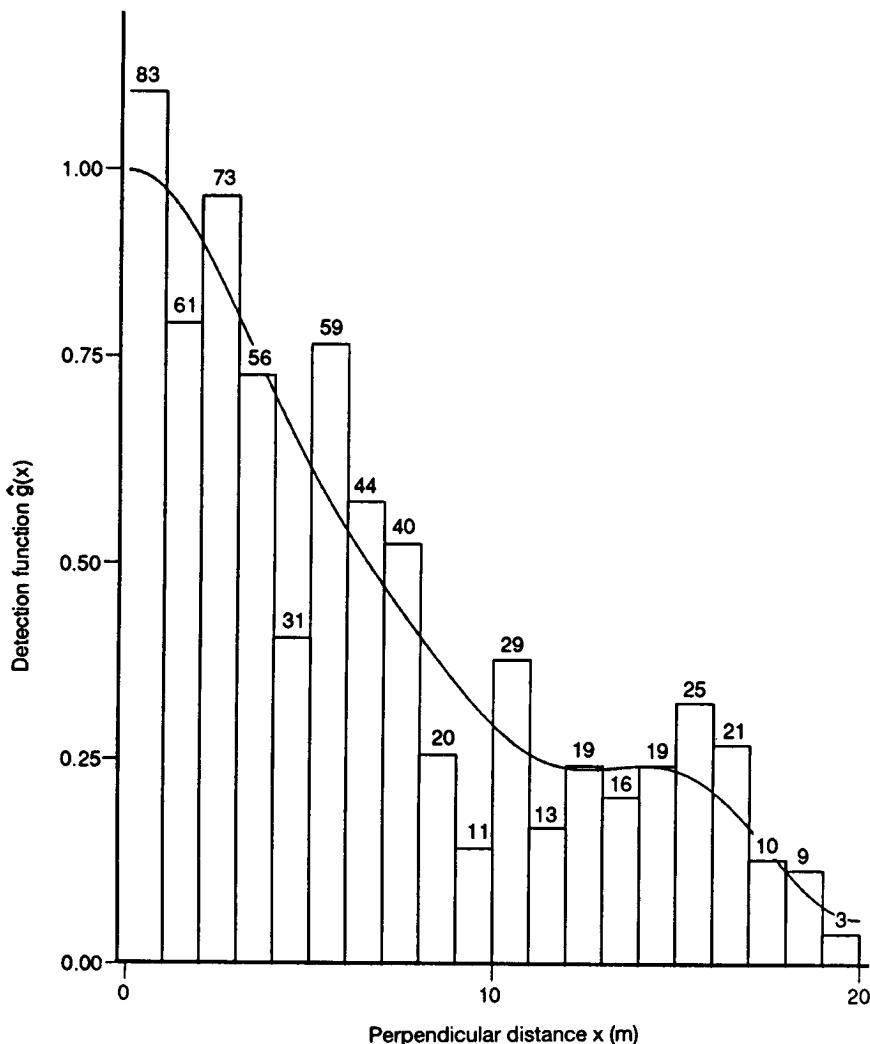
### WOODEN STAKE DATA

val coverage cannot be accurately judged from only 11 replicates, but examination of the intervals in Table 8.3 shows no particular indication of poor coverage. Estimates of density are low in all cases, except for survey 4 for the half-normal estimator. Averaging the density estimates over the 11 surveys indicates a negative bias of approximately 13–17%. Pooling the data over the 11 surveys provides an approximate estimate of bias of about – 7 to – 8%. The main reason for the negative bias seemed to be some lack of model fit near the line. Examination of the histograms and estimated detection functions seems to indicate that models commonly fit poorly near zero distance. Some of the negative bias is due to measurement error for stakes near the line. The exact location of each stake was known, and errors in measurement could be assessed. For example, the information for three stakes (i.e. stake numbers 45, 93 and 103) is shown:

Stake no.	103	45	93
True distance (m)	0.92	5.03	14.96
Ave. distance ( $\bar{x}$ )	0.73	4.77	14.63
$\widehat{sd}(x)$	0.139	0.151	0.277
$100(\widehat{sd}(x)/\bar{x})$	18.9	3.2	1.9

This suggests the measurement error is largely due to improper determination of the centreline. Finally, the negative bias is partially the result of observers missing about 4% of the stakes in the first metre and 13% of the stakes in the first two metres. However, this is offset by the tendency to underestimate distances near the line. One observer was seen actually tripping over a stake on the centreline, but still the stake was not detected. Stakes do not move and do not respond to the observer; for field surveys of animals, the relative importance of the different assumptions may be very different. If this survey was to be repeated, we would enlarge the study area and lengthen the line such that  $E(n) \doteq 80$ . Also, observers would be shown the evidence that stakes near the centreline were occasionally missed, and that measurements were often in error, in the hope that these problems could be lessened.

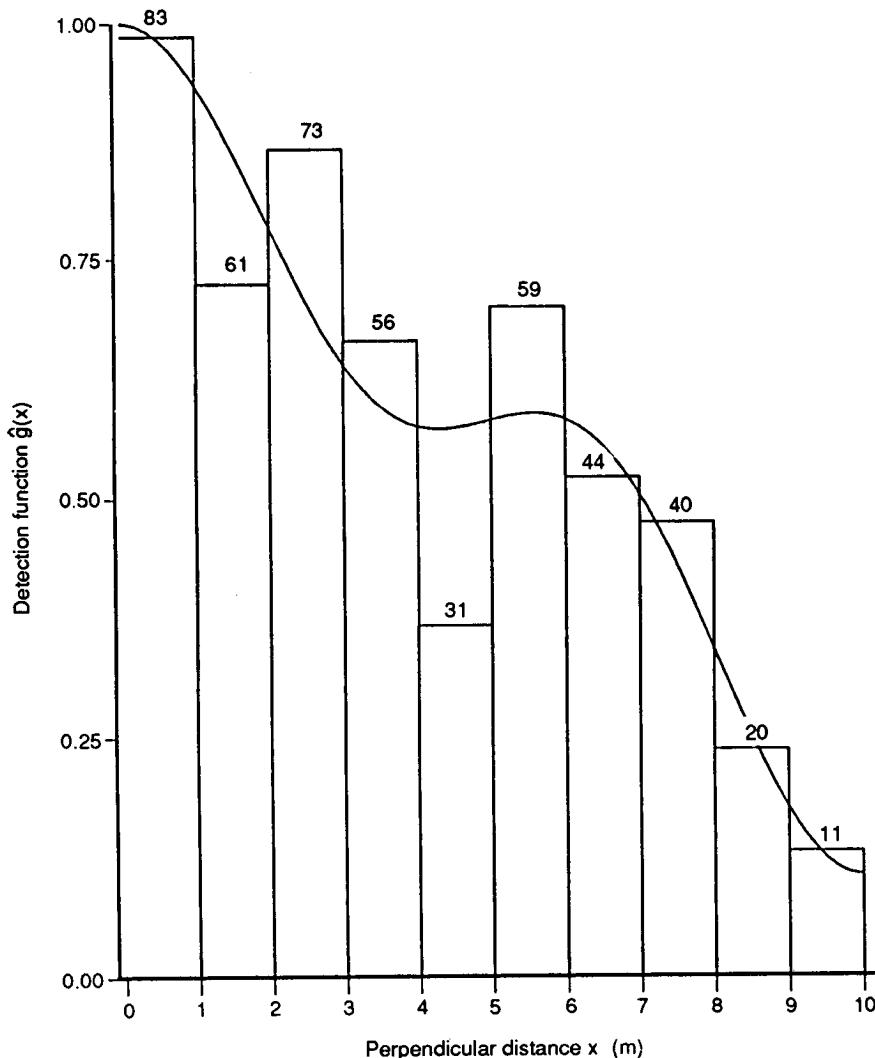
A second strategy will be illustrated that is less mechanical, requires a deeper level of understanding of the theory, and is somewhat more subjective. The pooled data are displayed as a histogram using 1 m cutpoints (Fig. 8.1). This information suggests several aspects that should influence the analysis. First, the distance data have a long tail, a small mode around 15 m, and considerable heaping at 10 m at the expense of 9 and 11 m. Some heaping at 0, 5 and 15 m is also seen. A histogram based on 0.5 m cutpoints indicated that detection declined rapidly near the line. Thus, the data have some of the characteristics



**Fig. 8.1.** Histogram of the wooden stake data of Laake (1978) using 1 m cutpoints. Also shown is a 3-term Fourier series fitted to these grouped data. Note the heaping at 0, 5 and 10 m, the relatively long tail in the distribution, and the additional mode near 15–16 m.

illustrated in Fig. 2.1. The modelling of  $g(x)$  will require additional terms to model the extra mode and long tail (Fig. 8.1 shows a model with three cosine terms, and this still seems to underestimate the data near the line). Thus, a reasonable approach might be to truncate the data at

### WOODEN STAKE DATA



**Fig. 8.2.** Histogram of the wooden stake data of Laake (1978), truncated at 10 m. The detection function is modelled using a uniform key function and three cosine adjustment terms.

$w = 10$  m in the hope of obtaining a better fit near the line and alleviating problems in the tail of the distribution.

At this point, one could choose a robust model for  $g(x)$  and proceed; here we will first use the uniform key and cosine adjustment function (Fourier series). Fitting these truncated data using a uniform key

## ILLUSTRATIVE EXAMPLES

function and a cosine adjustment function still required three terms, but provided a good fit near the line (Fig. 8.2). Some heaping at 5 m now shows more clearly. This model of the pooled data provided  $\hat{D} = 37.71$  ( $\widehat{\text{se}}(\hat{D}) + 2.82$ ) with a log-based confidence interval of (32.58, 43.64). These calculations assume  $\text{var}(n) = 0$ , as the stakes were placed uniformly. In reality, the value of  $\text{var}(n)$  in this study would be small but non-zero; zero was used only for illustration. Now, the data from each separate survey (i.e. individual person) can be modelled using a 3-term cosine series. The average estimate of density from the 11 surveys was 36.99 ( $\widehat{\text{sd}}(\hat{D}) = 5.04$ ), again close to the true density of 37.5.

Second, having selected a truncation point ( $w = 10$  m), one could select a model using AIC, based either on the individual data sets or on the pooled data sets. Here, we will examine only the pooled data. The results are summarized below.

<i>Model</i>	<i>AIC</i>
Half-normal + cosine	2118.7
Uniform + cosine	2112.4
Uniform + polynomial	2114.7
Half-normal + Hermite	2118.7
Hazard-rate + cosine	2126.8

The AIC selects the uniform + cosine model (the Fourier series) as the best model and, as shown above, provides an excellent estimate of density. The other models provided estimates that were inferior (half-normal + cosine = 33.60, uniform + polynomial = 31.47, half-normal + Hermite = 33.59, and hazard-rate + cosine = 31.90), and this could have been judged from the poor model fits near the line (Fig. 8.3). Likelihood ratio tests chose models with one parameter, except the hazard-rate model (two parameters) and the uniform + cosine (three parameters). Only the uniform + cosine model had confidence intervals that covered the true parameter.

Because the authors knew the value of  $D$  during these analyses, we cannot allege total objectivity in this example, which happened to provide an excellent estimate. However, the point is that careful review of the distance data can suggest anomalies (e.g. spiked data, long tails, heaping, a second mode), and these can suggest analysis approaches that should be considered equally with the more mechanical approach of using AIC or only likelihood ratio tests. We advocate some truncation, especially in cases where there are clear outliers in the distance data. Poor fit of the model to the data near the line should always be of concern. Generally the guidelines outlined in Section 2.5 will serve the analyst well in planning the analysis of distance data.

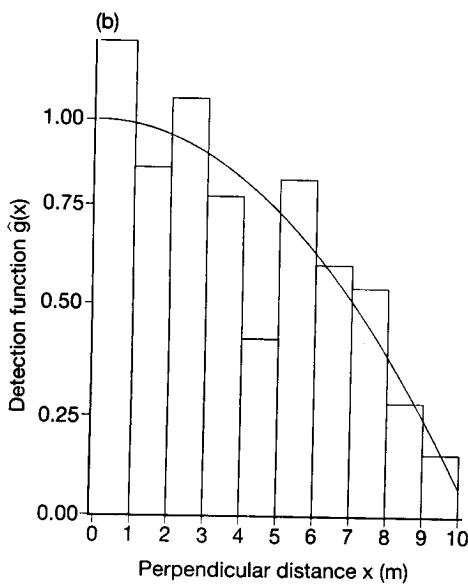
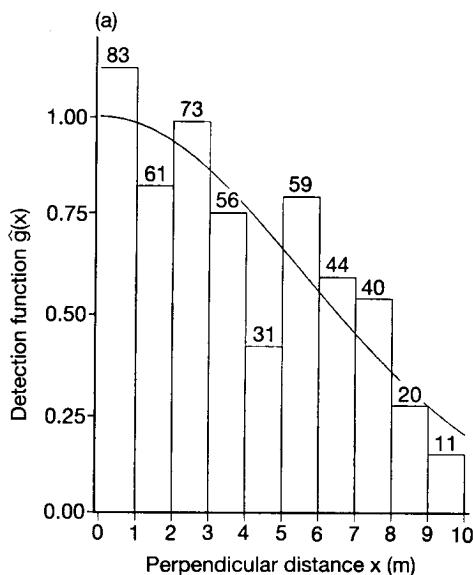
## 8.4 Studies of nest density

Studies of duck nest density have been conducted annually since 1964 at the Monte Vista National Wildlife Refuge in Colorado. During the 27 years of these studies, 10 041.6 miles of transect were walked and 4156 duck nests were found. Here we will examine the data for individual species and years to illustrate various points, approaches and difficulties. No attempt is made here to provide a final, comprehensive analysis of these data. Further details are found in Gilbert *et al.* (in prep.).

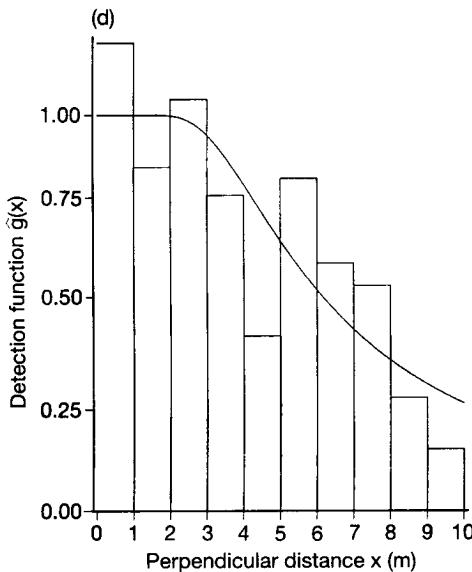
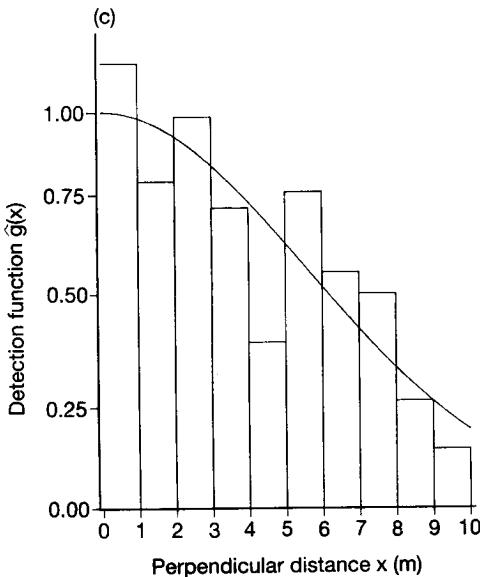
Strip transects were originally established on the refuge in a systematic design, running north-south, with 100 yds between transects and  $w = 8.25$  ft, giving a 5.5% sample of the entire 18.4 square mile refuge. Each transect centreline was marked by a series of numbered plywood signs attached to a 2.5 m pole (Burnham *et al.* 1980: 32). Only one-half of the original transects were surveyed during the 1969–90 period, except in 1971 when only one-quarter of the original transects were run and in 1977 when no survey was conducted. In 1969–79  $w$  was increased to 12 ft, but this was changed back to 8.25 ft during the 1980–85 period. Strip width was increased again to 12 ft in 1986–87 and finally changed back to 8.25 ft during 1988–90. Distances to detected nests were not recorded during 1964–66 and 1975–79. These erratic changes were often due to personnel or budget limitations. Transects were searched twice each year to monitor nest density of both early and later nesting species. A third search was made in a few years, but few nests were found and these data are not included in any of the examples given here. The mallard (*Anas platyrhynchos*) was the most common nesting duck, but substantial numbers of northern pintail (*Anas acuta*), gadwall (*Anas strepera*), northern shoveler (*Anas clypeata*), and teal (*Anas cyanoptera*, *A. discors*, and *A. carolinensis*) nests were also found. Species identity could not be determined for many nests; these were classed as ‘unknown’.

The refuge, at an elevation of 7500 ft, is characterized by level terrain and high desert vegetation in relatively simple communities. The drier, alkaline sites contain greasewood (*Sarcobatus vermiculatus*) and rabbit-brush (*Chrysothamnus* spp.) on the higher sites, while saltgrass (*Distichlis stricta*) dominates the lower sites. The wetter sites are dominated by baltic rush (*Juncus balticus*), but other species include cattail (*Typha latifolia*), spikerush (*Eleocharis macrosachya*), bullrush (*Scirpus validus*), and sedges (*Carex* spp.). Water is managed from pumped and artesian wells and irrigation sources and a system of dikes and borrow pits allow open water to be interspersed with vegetation cover to create good waterfowl nesting habitat. This area has one of the highest duck nest densities of any in North America.

## ILLUSTRATIVE EXAMPLES



## STUDIES OF NEST DENSITY



**Fig. 8.3.** Histograms of the wooden stake data (Laake 1978) and model fits for (a) the half-normal + cosine, (b) uniform + polynomial, (c) half-normal + Hermite, and (d) hazard-rate + cosine model.

## ILLUSTRATIVE EXAMPLES

The original design used strip transects, and it was assumed that all duck nests within the strip were detected. Perpendicular distance data were collected in 1967 and 1968, and it was clear by 1970 that some nests near  $w$  remained undetected, even with the narrow width of 8.25 ft (Anderson and Pospahala 1970). Perpendicular distances were recorded in 1969–74 and 1986–90, primarily as a means to relocate nests so that the nest fate could be determined. In years when  $w = 12$  ft, it was likely that more nests remained undetected. Thus, distance sampling theory is appropriate to obtain estimates of density, account for different sampling intensities, resolve differences in transect width  $w$ , and provide a basis for correction for undetected nests in years when no distance measurements were taken.

### **8.4.1 Spatial distribution of duck nests**

On biological grounds it seemed likely that nests were somewhat randomly distributed on the refuge. Thus, one might expect that the variation among the number of nests ( $n_i$ ) detected by transect line ( $l_i$ ) would be approximately Poisson, so that the variance in total sample size ( $n$ ) might be roughly Poisson (i.e.  $\widehat{\text{var}}(n) \doteq n$ ). The variance of  $n$  was computed empirically for mallard and non-mallard nests for each of 26 years using

$$\widehat{\text{var}}(n) = L \sum_{i=1}^k l_i \left( \frac{n_i}{l_i} - \frac{n}{L} \right)^2 / (k - 1)$$

where

$$L = \sum_{i=1}^k l_i$$

and

$$k = \text{number of replicate lines}$$

Modelling of  $\text{var}(n)$  as a function of  $n$  is common in statistical application (Carroll and Ruppert 1988). The ratio  $\hat{b} = \widehat{\text{var}}(n)/n$  was computed for each year for mallard and non-mallard nests (Table 8.4);  $b$  is often called a variance inflation factor. A random spatial distribution of nests yields  $\hat{b} \doteq 1.0$ . The relationship between  $n$  and  $\widehat{\text{var}}(n)$  can be estimated by a weighted linear regression through the origin, where the weight is the sample size. The point estimate under this approach is equivalent to a ratio estimator. The estimated value of  $b$  for mallards was 1.630

## STUDIES OF NEST DENSITY

**Table 8.4** Sample size ( $n$ ), estimates of the empirical variance of  $n$  ( $\widehat{\text{var}}(n)$ ), and their ratio ( $\hat{b}$ ) for mallard and non-mallard duck nests

Year	Mallards			Non-mallards		
	$n$	$\widehat{\text{var}}(n)$	$\hat{b}$	$n$	$\widehat{\text{var}}(n)$	$\hat{b}$
1964	142	210.20	1.48	119	201.27	1.69
1965	140	188.79	1.34	72	94.41	1.31
1966	151	238.01	1.58	109	146.69	1.35
1967	150	255.58	1.70	94	156.29	1.66
1968	176	281.84	1.60	139	240.67	1.73
1969	112	168.79	1.51	118	166.61	1.41
1970	103	166.44	1.62	126	225.92	1.79
1971	66	179.61	2.72	48	139.98	2.92
1972	63	99.10	1.57	64	102.08	1.60
1973	102	160.18	1.57	137	450.11	3.28
1974	69	157.18	2.28	40	54.43	1.36
1975	37	45.49	1.23	28	57.86	2.07
1976	49	61.89	1.26	46	72.48	1.58
1978	20	27.61	1.38	35	49.33	1.41
1979	13	9.73	0.74	40	96.06	2.40
1980	34	23.46	0.69	55	82.93	1.51
1981	50	105.63	2.11	54	79.27	1.47
1982	57	125.91	2.21	53	88.53	1.67
1983	79	123.10	1.56	35	48.69	1.39
1984	112	180.20	1.61	38	50.16	1.32
1985	70	112.88	1.61	82	105.49	1.29
1986	114	205.24	1.80	83	105.90	1.27
1987	130	164.00	1.26	132	195.82	1.48
1988	93	182.94	1.97	57	65.66	1.15
1989	74	103.67	1.40	56	58.63	1.05
1990	56	117.97	2.11	32	37.63	1.18
Wt. Ave.		1.630			1.677	

( $\widehat{s_e} = 0.068$ ) and 1.677 ( $\widehat{s_e} = 0.111$ ) for non-mallards. These estimates are not significantly different ( $z = 0.36$ ) and, thus, a pooled estimate of  $b$  for all species was computed as  $\hat{b} = 1.651$  ( $\widehat{s_e} = 0.063$ ). This weighted regression has an adjusted correlation of  $r^2 = 0.93$  (Fig. 8.4), and provides  $\widehat{\text{var}}(n) \doteq 1.7n$  which will be used in the remaining material for this example. This reflects some contagion in the distribution of duck nests, related, no doubt, to the variable distribution and quality of nesting habitat on the refuge. Minor species such as the redhead (*Aythya americana*), which nest in specialized habitat types, probably had a very non-random spatial distribution of nests (i.e.  $b >> 1.0$ ). Note that  $\text{var}(n)$  and  $b$  can be estimated for all years, even those where perpendicular distances were not recorded.

## ILLUSTRATIVE EXAMPLES

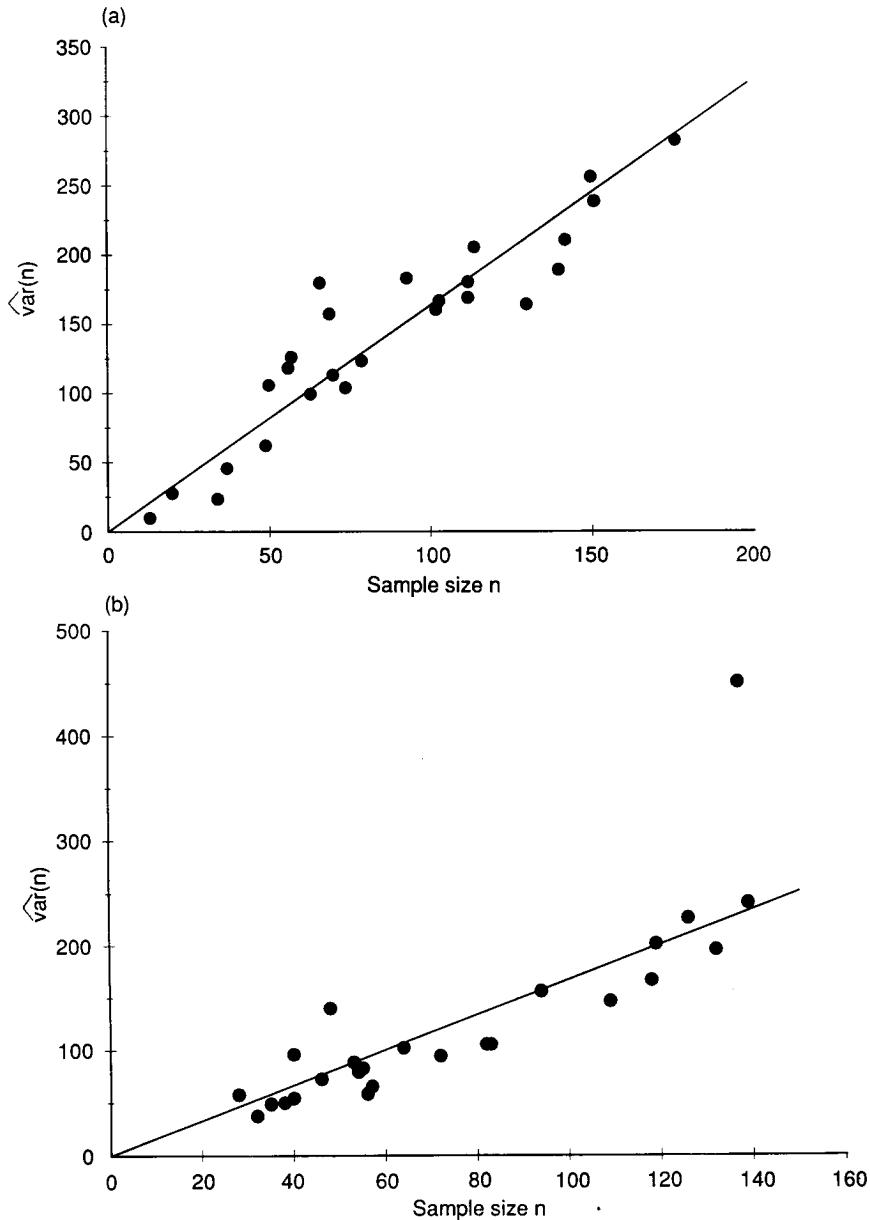
### 8.4.2 Estimation of density

This material focuses primarily on the grouped distance data collected during 1969–74 and 1986–87, years when  $w = 12$  ft. Histograms of the distance data are shown in Fig. 8.5 for the mallard, pintail, gadwall, teal, shoveler and unknown species. The first interval includes nests from 0 to 11 inches, the second interval includes nests detected from 12 to 23 inches, etc. Relatively few pintail nests were recorded within 2–3 ft of the centreline, whereas for other species, there appears to have been preferential heaping of nests close to the transect centreline. In each case, except the pintail, many nests found in the second interval (12–23 inches) were probably heaped into the first interval and perhaps into the third interval. Heaping at zero distance is common, and has been reported several times in distance sampling literature (e.g. Robinette *et al.* 1974), especially for data taken as sighting distance ( $r_i$ ) and angle ( $\theta_i$ ) and then the perpendicular distances ( $x_i$ ) computed as  $x_i = r_i \cdot \sin(\theta_i)$ . Also, there was a strong tendency to record nests somewhat outside the transect boundary ( $w$ ) as exactly 12 ft from the centreline. This latter form of heaping was right-truncated in all the analyses of the Monte Vista data in this chapter by setting  $w$  to 11.9 ft, thus excluding nests beyond 12 ft.

Estimated detection functions  $\hat{g}(x)$  are also plotted in Fig 8.5, assuming the half-normal key function with Hermite polynomial adjustments, if required. The half-normal key function seems quite reasonable for modelling these data. While there was substantial variation and some obvious heaping in the counts ( $n_i$ ), the fit appeared fairly good, with the clear exception of that for the pintail (discussed further in Section 8.4.3). Nests of the mallard seem to be most easily detected, as shown by the estimated unconditional probability of detection in the surveyed strip of area  $a = 2wL$  ( $\hat{P}_a = 0.80$ ,  $\hat{se} = 0.03$ ; Table 8.5). Gadwall nests were also easy to detect ( $\hat{P}_a = 0.76$ ,  $\hat{se} = 0.09$ ), whereas teal, shoveler, and unknown species nests were less detectable ( $\hat{P}_a = 0.64$ ,  $\hat{se} = 0.04$ ;  $\hat{P}_a = 0.60$ ,  $\hat{se} = 0.07$ ; and  $\hat{P}_a = 0.63$ ,  $\hat{se} = 0.04$ , respectively). Pooling all nests for all species results in an unconditional probability of detecting a nest within the transect of 0.78,  $\hat{se} = 0.02$ . Mallards (and pintails) nest early in the season when most vegetation has little new growth and detection might be easier. Other species tend to nest later and may experience more concealment in the vegetation.

Estimates of  $P_a$  for nests of all species combined were higher in years when  $w = 8.25$  ft ( $\hat{P}_a = 0.84$ ,  $\hat{se} = 0.07$ ) compared to years when  $w = 12$  ft ( $\hat{P}_a = 0.78$ ,  $\hat{se} = 0.02$ ). In general,  $P_a$  is a function of  $w$  in that  $P_a = 1/\{w \cdot f(0)\}$ ; under large sample approximations,  $cv(\hat{P}_a) = cv\{\hat{f}(0)\}$ . However, fewer nests were found using the narrower transect. The

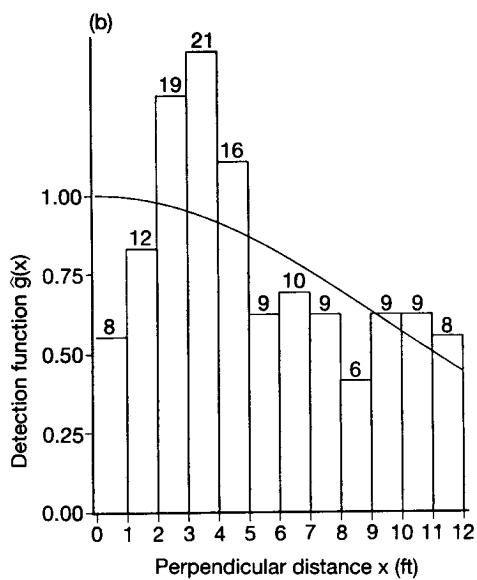
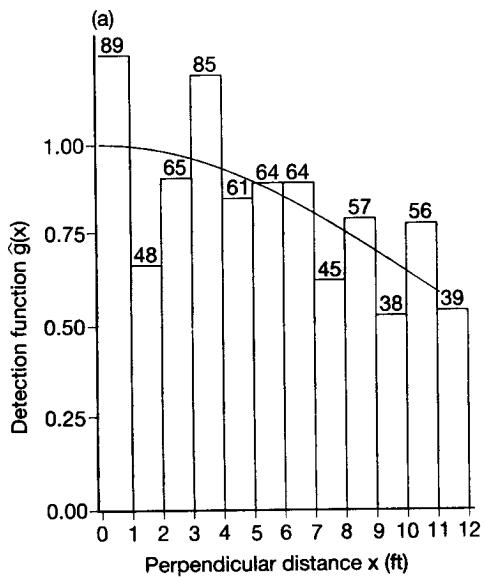
### STUDIES OF NEST DENSITY



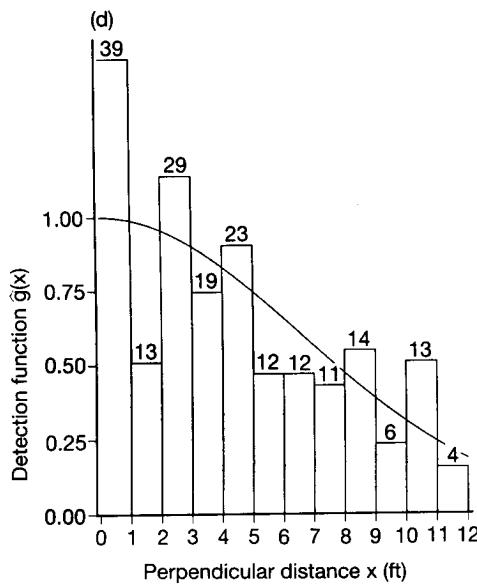
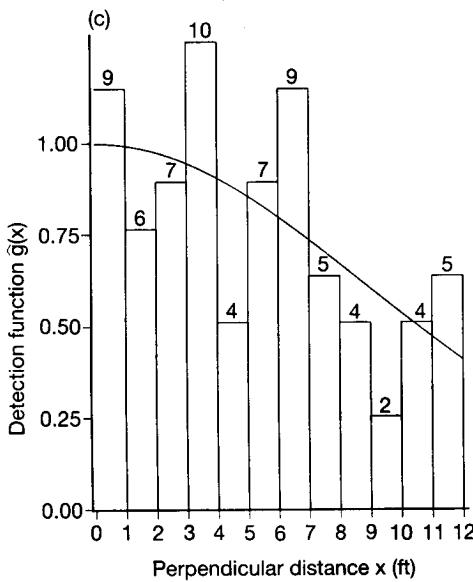
**Fig. 8.4.** Relationship between the empirical variance in  $n$  (i.e.  $\widehat{\text{var}}(n)$ ) and sample size ( $n$ ) for (a) mallard nests and (b) non-mallard nests.

narrow transects are inefficient as shown by the mean encounter rate (e.g. 0.867 and 1.137, respectively, for total nests). The values of  $w$

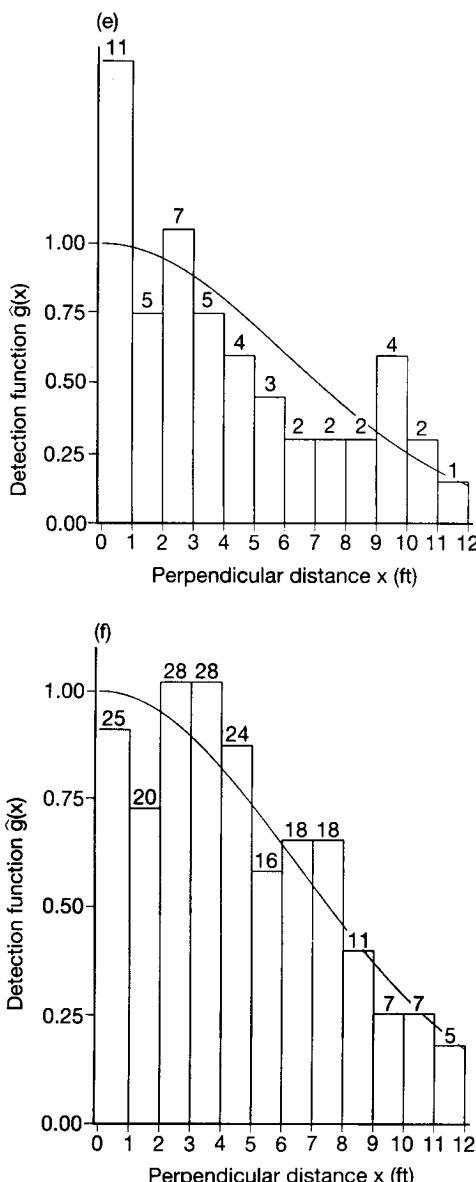
## ILLUSTRATIVE EXAMPLES



## STUDIES OF NEST DENSITY



## ILLUSTRATIVE EXAMPLES



**Fig. 8.5.** Histograms of the distance data for the nests detected at the Monte Vista National Wildlife Refuge in Colorado, USA, during 1969–74 and 1986–87. In addition, estimated detection functions using the half-normal key function and Hermite polynomial adjustments are shown for nests of (a) mallard ( $n = 711$ ), (b) pintail ( $n = 136$ ), (c) gadwall ( $n = 72$ ), (d) teal ( $n = 195$ ), (e) shoveler ( $n = 48$ ), and (f) unknown species ( $n = 207$ ).

## STUDIES OF NEST DENSITY

(either 8.25 or 12 ft) were certainly too large to meet the assumptions of strip transect sampling and too small for good efficiency in line transect sampling. Using the data for mallard nests, pooled for the 1969–74 and 1986–87 period,  $\hat{g}(8.25) \doteq 0.8$ , thus the observer undoubtedly detected many nests beyond  $w$  and could not record them. Detection near  $w$  is still good even for the wider transects [ $\hat{g}(12) \doteq 0.6$ ], thus many nests were still readily detectable beyond the transect boundary. As only about 0.58 mallard nests were found per transect mile (Table 8.5), it is clear that increasing  $w$  would allow sample size to increase with little additional survey effort. Finding a nest is a relatively rare event and if more nests could be found, little additional time would be required to take and record the relevant measurements. It would be interesting to increase  $w$  to perhaps 15, 18, or even 20 ft in future surveys to improve efficiency for both the observers and the estimators. An additional adjustment term might be required in modelling  $g(x)$  but an overall gain in the estimation process is likely. Observers would have to be cautioned to emphasize search on and near the centreline and not divert too much attention near  $w$ . Of course, it would be advantageous if heaping could be lessened and more accurate measurements were taken. Perpendicular distances might be remeasured as a test of accuracy when the fate of the nests is checked at a later time.

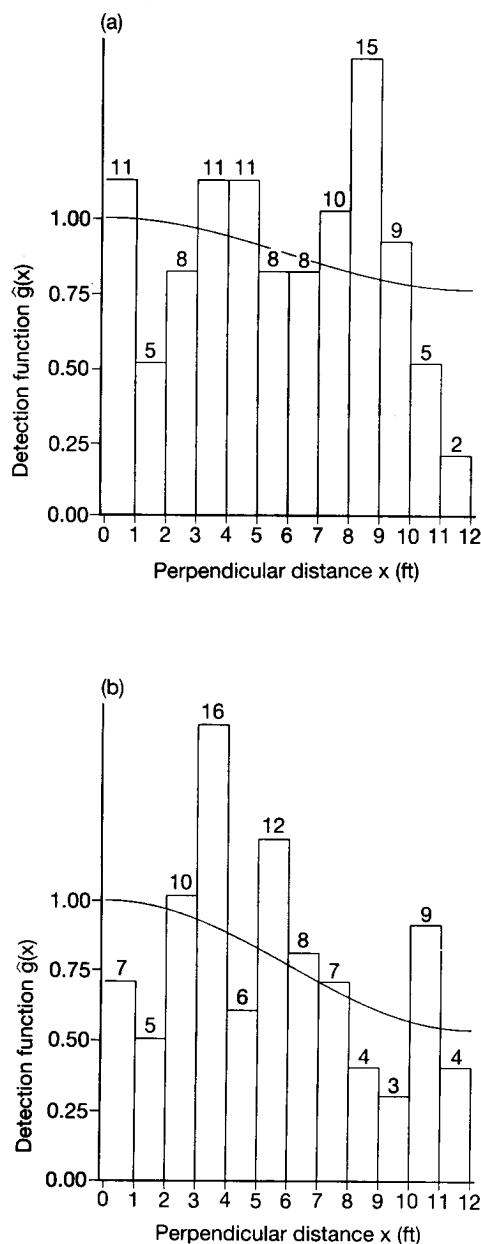
**Table 8.5** Summary of statistics and estimates for survey of duck nests at the Monte Vista National Wildlife Refuge in Colorado during 1969–74 and 1986–87. See Fig. 8.5 for histograms of these data and estimated detection functions using the half-normal model and Hermite polynomial adjustment terms. Density is in nests/mile<sup>2</sup>

Species	$n$	$\hat{P}_a$	$n/L$	$\hat{f}(0)$	$cv\{\hat{f}(0)\}$ (%)	$\hat{D}$	$cv(\hat{D})$ (%)*
All duck	1415	0.77	1.137	0.1079	2.6	323.8	3.7
Mallard	711	0.85	0.580	0.0901	3.8	149.6	5.4
Teal	195	0.64	0.157	0.1305	6.4	54.0	9.6
Pintail	136	0.78	0.109	0.1063	8.5	30.7	12.1
Gadwall	72	0.76	0.058	0.1094	11.5	16.7	16.5
Shoveler	48	0.60	0.039	0.1388	12.4	14.1	19.0
Unknown	207	0.63	0.166	0.1325	6.1	58.2	9.3

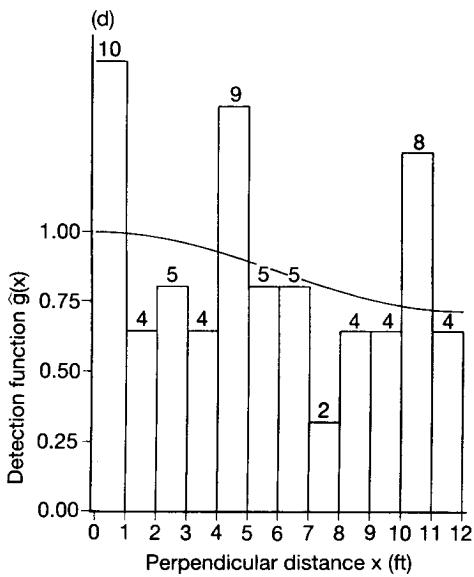
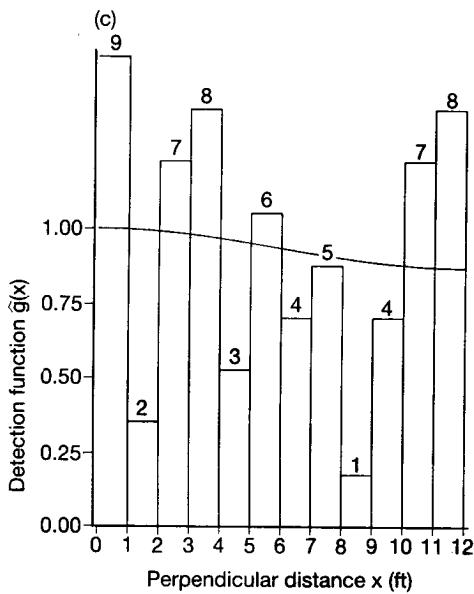
\* Assuming  $\widehat{\text{var}}(n) = 1.7 n$ .

Mallard nests were found in adequate numbers to allow annual estimates of nest density to be made. Histograms of the distance data are shown in Fig. 8.6, with estimated detection functions  $\hat{g}(x)$ . Although the annual sample sizes were generally fairly adequate ( $n_i > 60$  in every

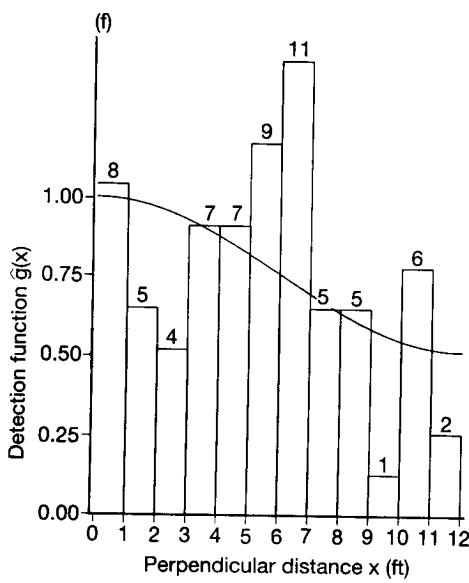
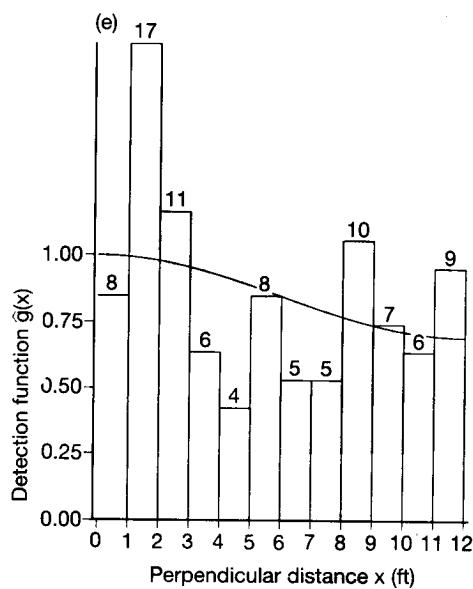
## ILLUSTRATIVE EXAMPLES



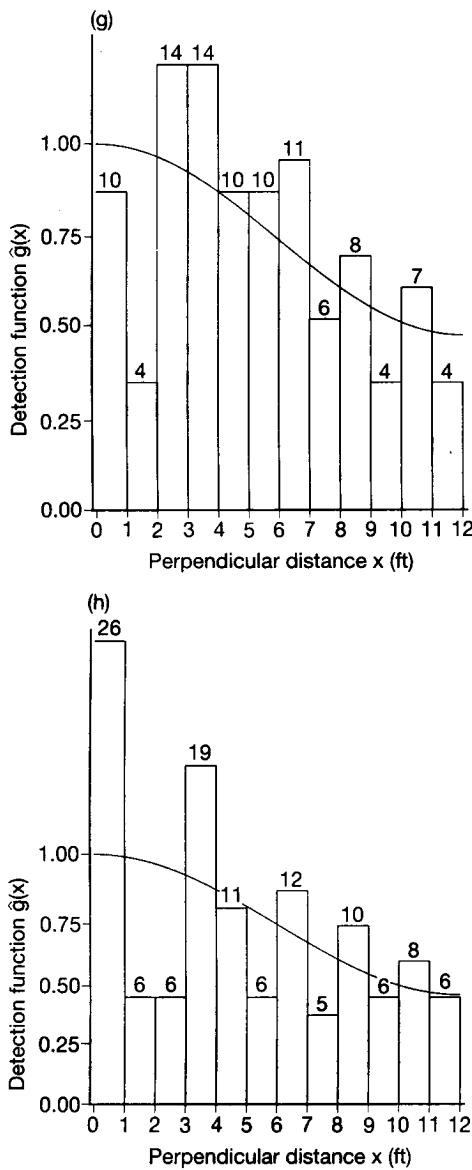
# STUDIES OF NEST DENSITY



## ILLUSTRATIVE EXAMPLES



STUDIES OF NEST DENSITY



**Fig. 8.6.** Histograms of the annual distance data for mallard nests detected at the Monte Vista National Wildlife Refuge in Colorado, USA, during 1969–74 and 1986–87. The estimated detection function is also shown. The respective sample sizes ( $n$ ) are (a) 103, (b) 91, (c) 64, (d) 64, (e) 96, (f) 70, (g) 102 and (h) 121.

## ILLUSTRATIVE EXAMPLES

year), the histograms look ‘rough’ and certainly exhibit heaping and perhaps some careless measurements. Except in the early years of the survey, observers generally had little training and were possibly not that motivated.

Estimation of annual nest density from data such as shown in Fig. 8.6 can be achieved using

$$\hat{D}_i = n_i \cdot \hat{f}_i(0)/2L_i, \quad \text{where } i = \text{year}$$

An attractive alternative exists if one is willing to make the assumption that the form of  $g_i(x)$  or  $f_i(x)$  is the same, or nearly the same, each year. This assumption seems biologically reasonable as the vegetation is low and somewhat sparse, nests of the various species do not move in response to the observer, the transects are rather narrow, and nests appear somewhat alike (but may vary somewhat depending on the stage of incubation). These reasons seem fairly compelling, but it is advisable to test the null hypothesis that the grouped distance data arose from a common detection function. Here a reasonable strategy of analysis is to fit the distance data, pooled over the eight years, to several good candidate models and select the model with the smallest AIC. These values for the mallard data under five models were:

<i>Model</i>	<i>AIC</i>
Half-normal + cosine	3513.8
Uniform + cosine	3513.9
Uniform + polynomial	3514.3
Half-normal + Hermite	3513.8
Hazard-rate + cosine	3515.4

Any of the five models could be used in this case, with little difference among the models. AIC selects the single-parameter half-normal key function with no adjustment terms, although the uniform key function plus a cosine term (i.e. Fourier series) is second best by a trivial margin. The uniform + polynomial model might also be a satisfactory model, followed by the 2-term hazard-rate model. For this example, the uniform + cosine (Fourier series) model will be used. Then, the eight individual distance data sets are analysed using the same model (i.e. the number of adjustment terms fixed) (Fig. 8.6), to give the log-likelihood value for the model based on the pooled data ( $\log_e(\mathcal{L}_p)$ ) and the eight log-likelihood values for the individual data sets ( $\log_e(\mathcal{L}_i)$ ). A likelihood ratio test may now be used:

$$\chi^2 = 2 \left[ \sum_{i=1}^8 \log_e(\mathcal{L}_i) - \log_e(\mathcal{L}_p) \right]$$

## STUDIES OF NEST DENSITY

where the test statistic is asymptotically distributed as  $\chi^2$  with  $n \cdot (r - 1)$  degrees of freedom ( $n$  = number of model parameters and  $r$  = number of data sets). With the mallard data,

$$\begin{aligned}\chi^2 &= 2[-1753.64 + 1755.93] \\ &= 4.58\end{aligned}$$

with  $1(8 - 1) = 7$  df, giving a  $p$ -value of 0.711. Thus, use of a common  $\hat{f}(x)$  for all years seems appropriate. One can pool the distance data over the eight years to obtain  $\hat{f}(0)$  and its standard error. Then, yearly estimates of density can be computed using the yearly sample size ( $n_i$ ) and

$$\hat{D}_i = n_i \cdot \hat{f}(0)/2L_i, \quad \text{where } i = \text{year}$$

The estimate of  $f(0)$  is specific to a given transect width, i.e. the value of  $\hat{f}(0)$  for data from a transect with  $w = 12$  ft cannot be used for years when the transect width was 8.25 ft. This approach seems appropriate for the mallard data, and it is nearly essential for species such as teal where yearly sample size would not support a reliable estimate of the year-specific detection function  $g_i(x)$  (or, equivalently,  $f_i(x)$ ). Data for the shoveler ( $n = 48$ ) and gadwall ( $n = 72$ ) support only an estimate of an average  $f(0)$  over the eight-year period, but this analysis approach allows annual estimates of density by using the year-specific sample sizes  $n_i$ . This general approach can be used for the analysis of other years of data where different transect widths were used.

Estimates of density under the five models were quite similar, ranging from 149.3 to 157.0 mallard nests per square mile. This might have been expected because the AIC values were all of similar magnitude. Models for  $g(x)$  contained only one parameter (either a key or an adjustment parameter), except the hazard-rate model, with two parameters. Thus, two of the five models used only the half-normal key function.

### **8.4.3 Nest detection in differing habitat types**

Despite large sample sizes, well distributed in time and space, the ability to examine differences in detectability by vegetation type was limited by the fact that baltic rush and greasewood made up approximately 68% and 15% of the vegetation on the refuge, respectively. Initially it was hypothesized that nest detectability would decline more rapidly with distance from the centreline in the tall, but often sparse, stands of greasewood when compared to the lower, more dense areas of rush. Instead, it became clear that the histogram of grouped distance data for nests found in greasewood indicated a mode well away from the transect centreline. It was hypothesized that observers would avoid the thorny

## ILLUSTRATIVE EXAMPLES



**Fig. 8.7.** Stand of greasewood with extensive areas of bare ground typify many upland sites on the Monte Vista National Wildlife Refuge. Observers may tend to avoid walking through these thorny shrubs, thus biasing the data.

greasewood (see Fig. 8.7) by walking off line and around these shrubs. Thus, nests at the base of these shrubs tended to go undetected near the transect centreline. Nests detected at the edge of greasewood clumps would be detected with near certainty while the observer was temporarily off the centreline (and thus avoiding the greasewood). Once such a nest

## STUDIES OF NEST DENSITY

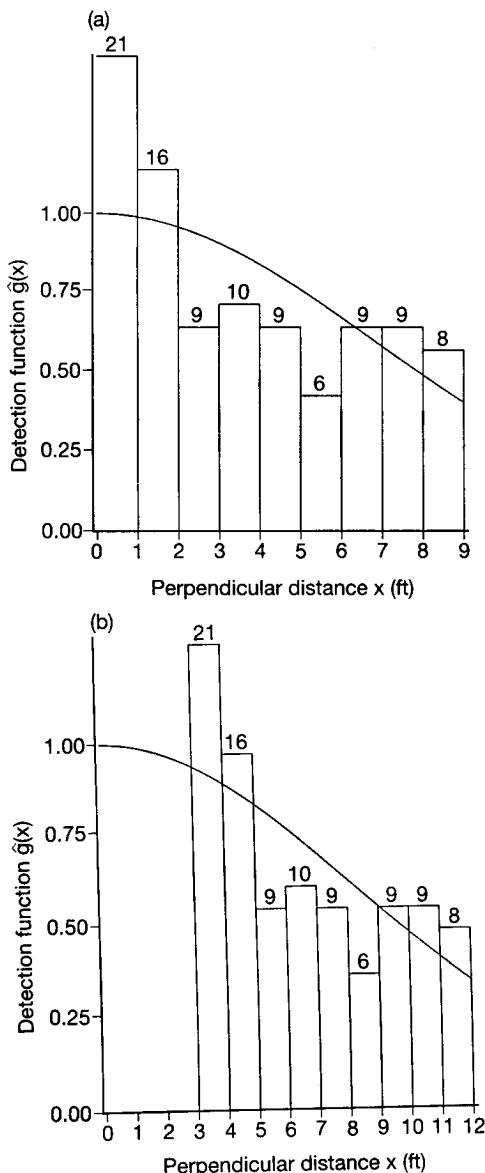
was found its distance to the centreline was measured and recorded. Such temporary departures from the transect centreline could explain the odd distance data for the pintail nests (Fig. 8.5). Perhaps pintail were common nesters in greasewood types and, thus, many were missed near the centreline. Indeed, 24.2% of the pintail nests were found in greasewood; surely this percentage would be still higher if nests near the centreline in greasewood were all detected. Other species nested in greasewood types less frequently: mallard 15.6%, gadwall 19.5%, teal 6.9%, and shoveler 2.6%. We tentatively conclude that observers were reluctant to enter the thorny greasewood type, and this resulted in nests being missed near the centreline.

An alternative explanation is that the observer measured the distance from his or her position to the nest and that pintail tended to nest at least two feet into the greasewood type. Then Fig. 8.5b would arise without missing any nests near the centreline; instead, the data would arise because the observer's path would go through habitat with a low pintail nest density. In any event, the presence of obstacles such as greasewood on the line must be dealt with effectively in the field survey or the analysis of the data can be problematic. We do not always advocate that the observer plunge through such cover types; instead, extra care in searching must be taken when an easier path is temporarily followed. For example, the observer could go around clumps of such vegetation both to the left and then to the right, searching the centreline more carefully. In any event, the measurements must be taken from the transect centreline, not to the observer who may be away from the centreline.

A definitive analysis of data such as those for the pintail nests is not possible. Approximate analyses that might be useful could be considered. First, one could fit a monotonically constrained function for  $g(x)$  as is shown in Fig. 8.5b for the half-normal key function with Hermite polynomial adjustments. This is likely to result in an underestimate of density if a substantial number of nests near the centreline was undetected. However, in this particular case, one knows from several other, similar species in this survey that the shape of  $g(x)$  has a broad shoulder, so that the procedure might be acceptable.

Second, one could use some arbitrary left-truncation and then estimate  $f(0)$  and  $D$  using, for example, the uniform + cosine or half-normal + Hermite model. First, one could decide on a truncation point; 3 ft might be reasonable for the pintail nest data. Here the grouped distance data less than 3 ft could be discarded, the remaining data rescaled as if the third interval was actually the first interval, and proceed to estimate density in the usual way (Fig. 8.8a). This is likely to be similar to the first procedure because we have reason to suspect that the detection function for pintail nests is fairly flat. Still, in this case, some underestimation

## ILLUSTRATIVE EXAMPLES



**Fig. 8.8.** Histograms of the distance data for pintail nests detected at the Monte Vista National Wildlife Refuge in Colorado, USA, during 1969–74 and 1986–87. Two estimates of  $g(x)$  are shown using alternative ways to left-truncate the data and minimize the problems observed in the first 3–4 distance categories. Left-truncation and rescaling are shown in (a), and the Alldredge and Gates (1985) approach is shown in (b)

## STUDIES OF NEST DENSITY

might be expected (unless  $g(0) \doteq 1.0$  but nests close to zero tended to be recorded at around 3 ft; then overestimation might result).

Third, the left truncation procedure of Alldredge and Gates (1985) could be employed, using the same truncation point. The result of this procedure is very dependent upon the model chosen and is often imprecise (Fig. 8.8b). In this example, where something is known about the distribution of distances of nests of other species of ducks, it seems likely that density of pintail nests is overestimated using this approach. Of course, any left truncation decreases sample size. The results of using the three approaches for the pintail nest data are summarized below for the half-normal key function and Hermite polynomial adjustments:

Method	<i>n</i>	$\hat{D}$	cv(%)
Full data	136	30.7	12.1
Left-truncate and rescale	97	29.8	14.2
Full left-truncation	97	35.0	17.9

The three estimates seem fairly reasonable for the pintail nest data, although one might prefer a density estimate near 30–32, rather than 35, unless the observer's path around greasewood types tended to sample areas of low pintail nest density. Considerable precision is lost in efforts to alleviate this problem; this is to be expected given the uncertainty introduced.

### 8.4.4 Models for the detection function $g(x)$

Various combinations of the key and adjustment functions provide flexibility in modelling the detection function  $g(x)$ . For data sets exhibiting a reasonable shoulder and meeting the other assumptions of

**Table 8.6** Summary of density estimates (above) and coefficients of variation (below) for five models of  $g(x)$  and four duck species, 1969–74 and 1986–87

Key function	Adjustment function	Gadwall ( <i>n</i> = 72)	Teal ( <i>n</i> = 195)	Shoveler ( <i>n</i> = 48)	Mallard ( <i>n</i> = 711)
Half-normal	Cosine	16.7	63.1	19.0	149.6
		16.5	12.4	22.4	5.4
Half-normal	Hermite	16.7	54.0	14.1	149.6
		16.5	9.6	19.0	5.4
Uniform	Cosine	17.4	54.8	14.0	155.2
		16.5	9.0	17.7	5.7
Uniform	Polynomial	15.9	61.5	12.2	147.3
		15.0	10.9	16.2	5.0
Hazard-rate	Cosine	16.2	67.0	17.9	147.6
		19.4	18.0	23.1	6.6

## ILLUSTRATIVE EXAMPLES

distance sampling, the choice of model, among those recommended here, is relatively unimportant. Estimates of density and estimated standard errors are summarized in Table 8.6 for several reasonable models for nest data on mallard, gadwall, teal and shoveler for the Monte Vista data. The differences in estimated density are small relative to the estimated standard errors. The standard errors given relate to the estimates made from data pooled over eight years.

If the distance data are distributed in a more spiked form, the choice of model is more difficult and the estimate of density more tenuous. The models recommended here are likely to perform reasonably well, except in pathological cases. A model with an appreciably smaller  $\chi^2$  goodness of fit value, if constrained to be non-increasing, will tend to be better than other models with the same number of estimated parameters. However, in general, goodness of fit tests are of relatively little help in model selection. In particular, some lack of fit near  $w$  is of little consequence in comparing model fit among several models.

### 8.5 Fin whale abundance in the North Atlantic

Large-scale line transect surveys of the North Atlantic to assess whale abundance were carried out in 1987 and 1989 (North Atlantic Sightings Surveys, NASS-87 and NASS-89). We analyse here the fin whale (*Balaenoptera physalus*) data from the 1989 survey collected by Icelandic vessels to illustrate the use of stratification. The analyses are extracted from Buckland *et al.* (1992b).

In 1989 four Icelandic vessels surveyed Icelandic and adjacent waters during July and August. The area covered was mostly within the East Greenland/Iceland stock boundaries for fin whales, and we consider here abundance estimation for that stock alone.

Sighting distances and angles were smeared and assigned to perpendicular distance intervals, using smearing method 2 of Buckland and Anganuzzi (1988a), and the hazard-rate model was fitted to the group frequencies. Detections were often of more than one animal, so an analysis of clusters was carried out; average cluster (school) size was roughly 1.5 whales. Several potential stratification factors were identified: geographic block, Beaufort (a scale for wind speed, generally determined from sea state), cloud cover, vessel and school size. Ideally stratification should be by all of these factors, but sample size considerations preclude this. Variables Beaufort, cloud cover and school size could be entered as covariates to avoid sample size difficulties, although it is then necessary to define a linear or generalized linear model between these effects and say effective strip width or encounter rate, and con-

founding between say Beaufort and geographic location, and hence between Beaufort and whale density, is inevitable. For analysing minke whale data, Gunnlaugsson and Sigurjónsson (1990) used generalized linear modelling to estimate sighting efficiency in different Beaufort states during NASS-87. This approach also has shortcomings when Beaufort varies strongly with geographic location, if whale density also varies geographically. For example, encounter rate may be lower in high Beaufort simply because a disproportionate amount of rough weather encountered by survey vessels was in an area with low animal density. Geographic stratification reduces but does not eliminate this effect. The problem of estimating fin whale abundance is easier than that of estimating North Atlantic minke whale abundance since cues are more visible. We adopt a simpler approach here to determine stratification factors.

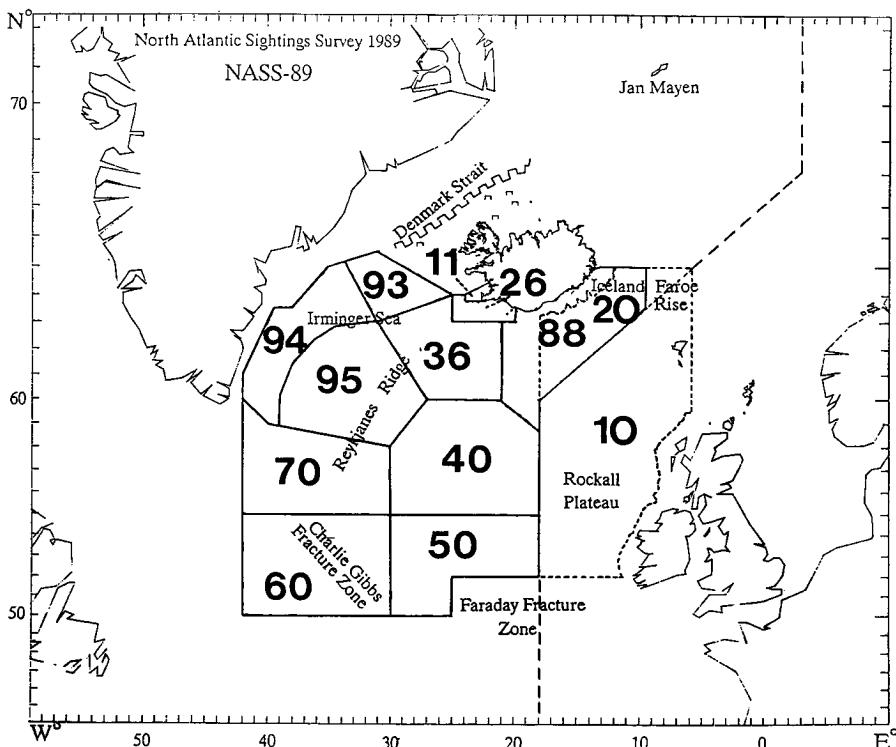
To assess say the effect of Beaufort, average school size, encounter rate and effective strip width were estimated for each Beaufort (0–6) in turn, pooling across all other possible stratification factors. Standard errors were calculated for each estimate, and z-tests carried out to assess whether there are significant differences in estimates at different Beauforts. Standard error for school size was calculated as sample standard deviation divided by square root of sample size; for encounter rate, the rate per day was calculated, and the sample variance of these rates, weighted by daily effort, used as described for the empirical method of Section 3.7.2; and the standard error for effective strip width was obtained from likelihood methods, *via* the information matrix. The stratification factors are confounded with each other, and the above approach ignores interactions between them; analyses are supplemented here by knowledge of likely effects of the different factors on the three components of estimation to determine an appropriate analysis. Thus results from z-tests are not used blindly; if a pairwise test indicates that effective strip width is wider at Beaufort 4 than Beaufort 1, it would be considered spurious, because it is counter to the knowledge that detection is easier in low Beaufort, whereas if there was a trend towards narrower effective strip widths as Beaufort increases, stratification would be deemed necessary.

Suppose mean size of schools detected during Beaufort 0 is  $\bar{s}_0$ , and during Beaufort 1,  $\bar{s}_1$ . Denote their standard errors by  $\widehat{se}(\bar{s}_0)$  and  $\widehat{se}(\bar{s}_1)$  respectively. Then a z-test is carried out by calculating

$$z = \frac{\bar{s}_0 - \bar{s}_1}{\sqrt{[\{\widehat{se}(\bar{s}_0)\}^2 + \{\widehat{se}(\bar{s}_1)\}^2]}}$$

The distribution of  $z$  is approximately normal. Thus if  $z > 1.96$  or  $z < -1.96$ , the mean school sizes differ significantly at the 5% level

## ILLUSTRATIVE EXAMPLES



**Fig. 8.9.** Geographic blocks for which abundance of the East Greenland/Iceland stock of fin whales is estimated from Icelandic 1989 data.

( $p < 0.05$ ). Evidence against the null hypothesis that the mean school size is the same in both sea states is strong if  $p$  is small, whereas if  $p$  is large, the data are consistent with the null hypothesis.

No significant differences in encounter rates by sea state (Beaufort 0–6) were found (Table 8.7). Mean school size did not differ significantly for Beauforts 0–3 or for Beauforts 4–6, but there was strong evidence that the mean of recorded school sizes in Beaufort  $\geq 4$  is smaller than for Beaufort 2 or 3. The effective strip width was significantly smaller for Beaufort 0 than for all other Beauforts except 5, and significantly larger for Beaufort 1 than for Beauforts 3, 4, 5 or 6 ( $p < 0.05$ ). No other differences were significant at the 5% level, although the effective strip width was significantly smaller at Beaufort 5 than at Beauforts 2, 3 and 6 at the 10% level. The unexpected result for Beaufort 0 corresponds to a very small sample size (13); otherwise, there is an indication that effective strip width decreases with Beaufort, which is what we would

FIN WHALE ABUNDANCE IN THE NORTH ATLANTIC

expect. We estimate densities separately by low Beaufort (0–3) and high Beaufort (4–6), and average resulting estimates across Beaufort categories, weighting by effort. This analysis is valid provided the probability of detection on the centreline,  $g(0)$ , is unity for both Beaufort categories; the effective strip width need not be the same for both categories.

A similar analysis of cloud cover produced no significant differences, except that the encounter rate at cloud cover 3 was significantly higher than at cloud cover 2 ( $p \doteq 0.01$ ), probably because relatively more cloud cover 3 occurred in areas of high fin whale density. If cloud cover 3 did increase detectability, effective strip width might be expected to increase, yet no pairwise comparisons provided any evidence of this ( $p > 0.2$  for all six pairwise tests).

**Table 8.7** Number of sightings (after truncation but before smearing), effective strip width, encounter rate and mean school size by sea state, Icelandic fin whale data, NASS-89. Standard errors in parentheses. Values in the same column with different superscript letters differ significantly ( $p < 0.05$ )

Beaufort	Number of sightings $n$	Effective strip width (n.m.)	Encounter rate (schools/100 n.m.)	Mean school size $\bar{s}$
0	13	0.55 (0.05) <sup>a</sup>	3.08 (1.90) <sup>a</sup>	1.69 (0.24) <sup>ab</sup>
1	42	2.37 (0.22) <sup>b</sup>	3.47 (0.54) <sup>a</sup>	1.48 (0.15) <sup>ab</sup>
2	83	2.00 (0.23) <sup>bc</sup>	4.19 (1.12) <sup>a</sup>	1.54 (0.08) <sup>a</sup>
3	78	1.60 (0.20) <sup>c</sup>	4.13 (0.47) <sup>a</sup>	1.63 (0.11) <sup>a</sup>
4	44	1.17 (0.29) <sup>c</sup>	2.55 (0.38) <sup>a</sup>	1.25 (0.10) <sup>b</sup>
5	33	0.49 (0.19) <sup>ac</sup>	2.42 (0.92) <sup>a</sup>	1.21 (0.10) <sup>b</sup>
6	18	1.61 (0.19) <sup>c</sup>	2.51 (1.73) <sup>a</sup>	1.22 (0.14) <sup>b</sup>

**Table 8.8** Number of sightings (after truncation but before smearing), effective strip width, encounter rate and mean school size by area, Icelandic fin whale data, NASS-89. Standard errors in parentheses. Blocks 11 (no sightings) and 26 (one sighting) are ignored. Values in the same column with different superscript letters differ significantly ( $p < 0.05$ )

Block	Number of sightings $n$	Effective strip width (n.m.)	Encounter rate (schools/100 n.m.)	Mean school size $\bar{s}$
36	54	0.94 (0.51) <sup>abc</sup>	4.86 (1.31) <sup>ab</sup>	1.35 (0.10) <sup>ab</sup>
40	15	1.88 (0.14) <sup>a</sup>	1.30 (0.75) <sup>cd</sup>	1.13 (0.14) <sup>a</sup>
50	23	2.07 (0.35) <sup>a</sup>	2.03 (1.06) <sup>bcd</sup>	1.35 (0.13) <sup>ab</sup>
60	36	1.31 (0.32) <sup>abc</sup>	3.26 (1.40) <sup>abc</sup>	1.36 (0.10) <sup>ab</sup>
70	9	0.68 (0.18) <sup>c</sup>	1.18 (0.51) <sup>cd</sup>	1.11 (0.12) <sup>a</sup>
88	32	1.68 (0.44) <sup>ab</sup>	2.57 (0.45) <sup>bc</sup>	1.56 (0.13) <sup>bc</sup>
93	70	1.87 (0.18) <sup>a</sup>	16.39 (1.96) <sup>e</sup>	1.69 (0.12) <sup>c</sup>
94	66	1.14 (0.21) <sup>bc</sup>	7.84 (2.28) <sup>ae</sup>	1.53 (0.12) <sup>bc</sup>
95	5	0.75 (0.31) <sup>abc</sup>	0.24 (0.20) <sup>d</sup>	1.20 (0.22) <sup>abc</sup>

## ILLUSTRATIVE EXAMPLES

The geographic blocks defined for Icelandic surveys in 1989 are shown in Fig. 8.9. Highly significant differences between some blocks in encounter rate and mean school size are unsurprising, and we stratify by block for each of these components of estimation. There are also several pairwise comparisons between blocks that indicate significant differences between effective strip widths. Blocks 40, 50 and 93 yield wide estimated effective strip widths, whereas the estimates for blocks 70 and 95 are small (Table 8.8). Given adequate sample size, stratification could be by block, as for encounter rate and mean school size. However, effective strip width estimation is unreliable for small samples. There were only nine sightings in block 70 and five in block 95, rendering comparisons between them and other blocks of little value. Thus just the differences between block 94 and blocks 40 ( $p \doteq 0.04$ ), 50 ( $p \doteq 0.03$ ) and 93 ( $p \doteq 0.05$ ) are genuine cause for concern. For estimating effective strip width, we choose here to stratify the area into two parts: south (blocks 40, 50, 60 and 70) and north, since this also effectively stratifies by vessel type (below).

**Table 8.9** Number of sightings (after truncation but before smearing), effective strip width, encounter rate and mean school size by vessel, Icelandic fin whale data, NASS-89. Standard errors in parentheses. Values in the same column with different superscript letters differ significantly ( $p < 0.05$ )

Vessel	Number of sightings $n$	Effective strip width (n.m.)	Encounter rate (schools/100 n.m.)	Mean school size $\bar{s}$
<i>Sk</i>	43	1.09 (0.43) <sup>a</sup>	1.79 (0.47) <sup>a</sup>	1.26 (0.08) <sup>a</sup>
<i>AF</i>	49	1.15 (0.26) <sup>a</sup>	1.45 (0.21) <sup>a</sup>	1.31 (0.08) <sup>a</sup>
<i>Hv8</i>	83	1.43 (0.33) <sup>a</sup>	4.58 (0.47) <sup>b</sup>	1.43 (0.08) <sup>ab</sup>
<i>Hv9</i>	136	1.37 (0.17) <sup>a</sup>	8.02 (2.46) <sup>b</sup>	1.61 (0.08) <sup>b</sup>

The three components of estimation were also considered by vessel (Table 8.9). Most pairwise comparisons between vessels for encounter rate were significant, as were many of those for mean school size. These differences arise largely because vessels operated in different blocks; there is strong confounding between vessel differences and block differences. If vessel differences in encounter rate in particular occurred because different vessels have different searching efficiencies, significant differences in effective strip width between vessels might be anticipated, yet none were close to significance ( $p > 0.2$  in all pairwise tests). Given the similarity in effective strip widths across vessels, we pool distance data across vessels prior to analysis. Effective strip widths for the two research vessels (1.09 n.m. and 1.15 n.m.) were slightly smaller than for

## FIN WHALE ABUNDANCE IN THE NORTH ATLANTIC

the whaling vessels (1.43 n.m. and 1.37 n.m.). Although these differences are not significant, the impact on the analyses of estimating the effective strip width for research vessels separately from that for whaling vessels was assessed, and found to be slight. Because all effort in southern blocks was carried out by research vessels and most effort in the northern blocks was by whaling vessels, the decision to estimate the effective strip width separately for the northern and southern blocks, and to estimate encounter rate and mean school size by individual block, in effect gives stratification by vessel type.

Effective strip width did not show significant differences by size of school at the 5% level (Table 8.10), although there was a weak indication that the effective width was greater for schools of four or more animals than for single animals ( $p \doteq 0.1$ ). Since 68% of sightings were of single animals, and a further 22% were of pairs, the effect of variation in detectability due to school size on abundance estimates will be slight. However, stratification by school size is likely to be more valid and was adopted. Estimated effective strip width is almost identical for single animals and for pairs, and very few schools of more than three animals were detected, so two strata were defined: small schools (one or two animals) and large schools (three or more). Small sample sizes forced one modification to the preferred method of analysis: the number of large schools was too small to allow estimation of effective strip width separately for high and low Beaufort, so that for large schools only, a pooled estimate of effective strip width across Beaufort categories was calculated.

**Table 8.10** Number of sightings (after truncation but before smearing), effective strip width and encounter rate by school size, Icelandic fin whale data, NASS-89. Standard errors in parentheses. Values in the same column with different superscript letters differ significantly ( $p < 0.05$ )

School size	Number of sightings $n$	Effective strip width (n.m.)	Encounter rate (schools/100 n.m.)
1	211	1.27 (0.16) <sup>a</sup>	2.27 (0.16) <sup>a</sup>
2	68	1.25 (0.28) <sup>a</sup>	0.71 (0.13) <sup>b</sup>
3	22	1.50 (0.25) <sup>a</sup>	0.23 (0.05) <sup>c</sup>
4	8	1.71 (0.23) <sup>a</sup>	0.09 (0.03) <sup>d</sup>
> 4	2	—	—

To assess the impact of the decision to stratify by school size on estimates, two further analyses were carried out. The first of these

## ILLUSTRATIVE EXAMPLES

was exactly as above, except data were not stratified by school size. In the second, the data were reanalysed with individual animals as the sighting unit. Thus a school of size three between 0.75 and 1.0 n.m. perpendicular distance contributes a frequency count of three to that distance interval (before smearing). This method of analysis is used in the southern hemisphere minke whale subcommittee of the International Whaling Commission, so that school size can be estimated as the ratio of animal density, from this analysis, to school density, estimated conventionally. We do not recommend this approach in general, although it can be effective, if variances are estimated by robust methods. When estimates were summed across geographic blocks prior to combining across Beaufort categories, the preferred method of analysis (stratifying by school size) gave a total estimate of 11 054 whales ( $\widehat{se} = 1670$ ). Without stratification by school size, the estimate was 11 702 whales ( $\widehat{se} = 1896$ ). When individual whales were taken as the sampling unit, an estimate of 11 758 whales ( $\widehat{se} = 1736$ ) was obtained. Note that this latter strategy gave a very similar standard error to the other methods, even though sightings of individual whales were not independent events, and sample size is thus artificially increased. This occurs because of the robust method of estimating the variance in encounter rate, found by calculating the sample variance of the rate per day, weighted by daily effort, used as described for the empirical method of Section 3.7.2. Nevertheless, this approach underestimates the variance in effective strip width, unless it is obtained by resampling methods.

Abundance estimates for the East Greenland/Iceland stock of fin whales are given by block in Table 8.11. The sum of these estimates does not equal the corresponding estimate of 11 054 whales given above, due to the effects of calculating a weighted average of high and low Beaufort estimates within each block instead of first combining across blocks. The two estimates would be equal if the proportion of effort at low Beaufort was the same in every block. Suppose that 50% of effort occurred at low Beaufort overall, but in a given block, just 5% of effort occurred at low Beaufort. The method of summing estimates across blocks before averaging across Beaufort categories would give equal weight to the low and high Beaufort estimates in this block, whereas the method of Table 8.11 would give the high Beaufort estimate 19 times the weight of the low Beaufort estimate. The latter method is more appropriate, so the final abundance estimate of 10 378 whales ( $\widehat{se} = 1655$ ) is obtained by weighting the low and high Beaufort estimates by respective effort in individual blocks.

This example shows that reliable abundance estimates may be obtained by geographic block even when sample size within a block is very small.

## FIN WHALE ABUNDANCE IN THE NORTH ATLANTIC

**Table 8.11** Abundance estimates by block, East Greenland/Iceland fin whale stock, 1989

Block	Number of sightings $n$	$\hat{D}$ (whales/10 000 n.m. $^2$ )	$\widehat{\text{se}}(\hat{D})$	Size of block (n.m. $^2$ )	Abundance estimate $\hat{N}$	$\widehat{\text{se}}(\hat{N})$
36	54	270	72	44 172	1195	316
40	15	68	39	107 842	735	421
50	23	87	57	99 750	865	569
60	36	158	67	131 458	2071	879
70	9	74	58	88 571	658	517
88	32	129	38	59 848	770	230
93	70	873	220	21 761	1900	480
94	66	450	101	46 092	2073	467
95	5	16	14	69 396	111	95
All	323	155	25	668 891	10 378	1655

In two of the blocks of Table 8.11, sample size was under 10, yet analysis was possible stratifying not only by block but also by school size and Beaufort category. Of the three components of estimation, only effective strip width (or equivalently,  $f(0)$ ) cannot be reliably estimated when samples are small. If this parameter can be assumed to be constant across at least some of the stratification categories, small sample size problems are avoided. The method is far superior to prorating a total estimate, obtained by pooling data across blocks, between blocks according to their respective areas, which requires that density of animals is uniform across the entire surveyed area. Variance estimation requires some care, since the individual block estimates are not independent. Provided the common component ( $\hat{f}(0)$ ) of the respective estimates is removed when calculating the variance of their sum, then incorporated in the variance estimate using the delta method for approximating the variance of a product, as described in Section 3.7.1, valid variance estimates can be obtained quite simply. For the relatively complex fin whale analyses, variances are found as follows.

Within a stratum, abundance  $N$  is estimated by

$$\hat{N} = \frac{n \cdot \hat{f}(0) \cdot \bar{s} \cdot A}{2L}$$

with

$$\widehat{\text{var}}(\hat{N}) = \hat{N}^2 \cdot \left[ \frac{\widehat{\text{var}}(n)}{n^2} + \frac{\widehat{\text{var}}\{\hat{f}(0)\}}{\{\hat{f}(0)\}^2} + \frac{\widehat{\text{var}}(\bar{s})}{\bar{s}^2} \right]$$

## ILLUSTRATIVE EXAMPLES

where  $n$  = number of sightings within 3 n.m. of the centreline in the stratum,

$\hat{f}(0)$  = estimated probability density of perpendicular distances, evaluated at zero,

$\bar{s}$  = mean school size,

$L$  = distance covered while on effort,

$A$  = size of the area containing the population of  $N$  animals.

For a given block, the above yields independent estimates of  $f(0)$ , and hence of animal abundance corresponding to small schools in low Beaufort  $\hat{N}_{sm, lo}$ , small schools in high Beaufort  $\hat{N}_{sm, hi}$ , and large schools  $\hat{N}_{la}$  (unstratified by Beaufort). Then an estimate of abundance for animals in small schools is obtained by taking an average, weighted by effort carried out at low Beaufort ( $L_{lo}$ ) and high ( $L_{hi}$ ):

$$\hat{N}_{sm} = \frac{L_{lo} \cdot \hat{N}_{sm, lo} + L_{hi} \cdot \hat{N}_{sm, hi}}{L_{lo} + L_{hi}}$$

and

$$\widehat{\text{var}}(\hat{N}_{sm}) = \frac{L_{lo}^2 \cdot \widehat{\text{var}}(\hat{N}_{sm, lo}) + L_{hi}^2 \cdot \widehat{\text{var}}(\hat{N}_{sm, hi})}{(L_{lo} + L_{hi})^2}$$

An abundance estimate for the block is then

$$\hat{N}_{bl} = \hat{N}_{sm} + \hat{N}_{la}$$

with

$$\widehat{\text{var}}(\hat{N}_{bl}) = \widehat{\text{var}}(\hat{N}_{sm}) + \widehat{\text{var}}(\hat{N}_{la})$$

Within say the northern blocks, for which  $\hat{f}(0)$  estimates are in common, total abundance is estimated by

$$\hat{N}_N = \sum \hat{N}_{bl}$$

where summation is over all northern blocks. To estimate the variance of this estimate, note that  $\hat{N}_{bl}$  may be expressed as

$$\begin{aligned} \hat{N}_{bl} &= \frac{L_{lo} \cdot \hat{N}_{sm, lo} + L_{hi} \cdot \hat{N}_{sm, hi}}{L_{lo} + L_{hi}} + \hat{N}_{la} \\ &= \frac{L_{lo} \cdot \hat{f}_{sm, lo}(0) \cdot \hat{M}_{sm, lo} + L_{hi} \cdot \hat{f}_{sm, hi}(0) \cdot \hat{M}_{sm, hi}}{L_{lo} + L_{hi}} + \hat{f}_{la}(0) \cdot \hat{M}_{la} \\ &= l_{lo} \cdot \hat{f}_{sm, lo}(0) \cdot \hat{M}_{sm, lo} + l_{hi} \cdot \hat{f}_{sm, hi}(0) \cdot \hat{M}_{sm, hi} + \hat{f}_{la}(0) \cdot \hat{M}_{la} \end{aligned}$$

FIN WHALE ABUNDANCE IN THE NORTH ATLANTIC

where

$$l_{lo} = \frac{L_{lo}}{L_{lo} + L_{hi}}$$

$$l_{hi} = \frac{L_{hi}}{L_{lo} + L_{hi}}$$

$$\hat{M}_{sm, lo} = \frac{n_{sm, lo} \cdot \bar{s}_{sm, lo} \cdot A}{2L_{lo}}$$

evaluated for that block, and similarly for  $\hat{M}_{sm, hi}$  and  $\hat{M}_{la}$ . The component  $\hat{f}(0)$  is common across blocks, whereas the other components of the abundance estimate are not. Thus

$$\hat{N}_N = \hat{f}_{sm, lo}(0) \cdot \sum [l_{lo} \cdot \hat{M}_{sm, lo}] + \hat{f}_{sm, hi}(0) \cdot \sum [l_{hi} \cdot \hat{M}_{sm, hi}] + \hat{f}_{la}(0) \cdot \sum \hat{M}_{la}$$

where summation is over blocks. Denote the three terms in this expression by  $T_i$ ,  $i = 1, 2, 3$ ; these three terms are independent. Consider the final term,

$$T_3 = \hat{f}_{la}(0) \cdot \sum \hat{M}_{la}$$

This has variance

$$\widehat{\text{var}}(T_3) = T_3^2 \cdot \left[ \frac{\widehat{\text{var}}\{\hat{f}_{la}(0)\}}{\{\hat{f}_{la}(0)\}^2} + \frac{\sum \widehat{\text{var}}(\hat{M}_{la})}{\{\sum \hat{M}_{la}\}^2} \right]$$

where

$$\widehat{\text{var}}(\hat{M}_{la}) = \hat{M}_{la}^2 \cdot \left[ \frac{\widehat{\text{var}}(n_{la})}{n_{la}^2} + \frac{\widehat{\text{var}}(\bar{s}_{la})}{\bar{s}_{la}^2} \right] \text{ evaluated in each block}$$

Similarly,

$$\widehat{\text{var}}(T_1) = T_1^2 \cdot \left[ \frac{\widehat{\text{var}}\{\hat{f}_{sm, lo}(0)\}}{\{\hat{f}_{sm, lo}(0)\}^2} + \frac{\sum l_{lo}^2 \cdot \widehat{\text{var}}(\hat{M}_{sm, lo})}{\{\sum l_{lo} \cdot \hat{M}_{sm, lo}\}^2} \right]$$

and likewise for  $\widehat{\text{var}}(T_2)$

## ILLUSTRATIVE EXAMPLES

Finally,  $\widehat{\text{var}}(\hat{N}_N) = \sum_{i=1}^3 \widehat{\text{var}}(T_i)$

If total abundance in the southern blocks is estimated by  $\hat{N}_S$ , it and its variance are estimated in the same way as for  $\hat{N}_N$ . Total abundance over the whole area is then the sum of these estimates, with variance equal to the sum of the respective variances, since  $f(0)$  is estimated independently in the two areas. Applying the above methods, we obtain an abundance estimate of 10 378 fin whales, with standard error 1655. Assuming  $\hat{N}$  is lognormal, the estimated 95% confidence interval for  $N$  is (7607, 14 158) animals (Section 3.7.1).

## 8.6 Use of tuna vessel observer data to assess trends in abundance of dolphins

In some circumstances, the scientist has little control over the design of line transect surveys, so that robust analysis techniques must be used to cope with potentially biased data. An example is the extensive database gathered by observers placed on board tuna vessels in the eastern tropical Pacific. Large tuna, and in particular the yellowfin tuna (*Thunnus albacares*), associate with dolphin schools in this region, and fishermen use speed boats to herd dolphins into large purse-seine nets to catch the tuna under them. Although most dolphins are released again, high mortality, largely through entanglement, can occur. To estimate this mortality, and to assess its impact on dolphin stock size, the observers record data on many variables, including sighting angles and distances to detected schools. We use these data for one of the affected stocks, the northern offshore stock of spotted dolphin (*Stenella attenuata*), to illustrate the techniques of Buckland and Anganuzzi (1988b) and Anganuzzi and Buckland (1989), which attempt to estimate trends in abundance that are robust to the many biases inherent in sightings data from fishing vessels. This had been attempted earlier, notably by Hammond and Laake (1983).

The database comprises hundreds of cruises and thousands of detections of dolphin schools, which have accumulated annually since 1975. Thus, the analyst can concentrate on reducing the effects of biases in the data, to yield smoother trends in relative abundance estimates, at the expense of precision.

Equation 3.4 gave the general formula for estimating animal density from line transect data. For tuna vessel dolphin sightings data, the constant  $c = 1$ , and  $g_0$  may be assumed to be unity, since school sizes are large and the sighting cue is continuous. Thus density may be expressed as

## USE OF TUNA VESSEL OBSERVER DATA

$$D = \frac{E(n) \cdot f(0) \cdot E(s)}{2L}$$

If the stock area is of size  $A$ , the stock size may be expressed as

$$N = \frac{E(n)}{L} \cdot f(0) \cdot E(s) \cdot \frac{A}{2} \quad (8.1)$$

The first term in this expression is the encounter rate, the second term is the density function evaluated at zero, which is the reciprocal of the effective strip width, the third term is the true mean school size for the stock, and the remainder is a known constant. Thus there are three components to estimate. Because tuna vessel search effort is concentrated where captains expect to catch tuna, correction is required to adjust for search effort. Since the methods described here are to monitor changes in abundance, absolute abundance estimates are not required, but we must assume that bias in the relative abundance estimates is consistent between years, i.e. percentage relative bias should be constant. However, in a year of abundance of yellowfin tuna, effort is concentrated on dolphin schools, whereas in a year of scarcity, it becomes more economic to catch tuna too small to associate with dolphins. Thus if search effort is not corrected for, we cannot expect bias to be consistent. We correct for search effort in the following way.

Suppose for a random point in the ocean we can estimate the expectation of each of the three components, encounter rate,  $f(0)$  and school size. Then their product, multiplied by the stock area divided by two, provides an estimate of abundance. The estimates need not be independent, provided variance estimation allows for correlation. We can therefore reduce the estimation problem to three simpler problems, each comprising a single parameter. Note that it is the expectation of each component for a random point in the ocean that is required. Thus an estimate of the true mean school size of the stock is inappropriate, since in areas of high school density, average school size may also be high. Such estimation would not be problematic if each component showed very little variation through the stock area. In practice variation is large, but if strata are defined such that the variation in a component within each stratum is small relative to variation between strata, the component can be estimated within each stratum, and averaged across strata, weighting by the area of each stratum. A method of stratifying the area is therefore required. The stratification can be different for each component. (Another option is to stratify by search effort, and estimate all three components from this single stratification. However, the low effort

## ILLUSTRATIVE EXAMPLES

stratum proves to be very heterogeneous, and typically contains a high proportion of the stock, and the method lacks robustness.)

The stock area is post-stratified using the tuna vessel data. For each component, a variable that correlates well with that component must be identified. We use crude encounter rate, average school size and average sighting distance, calculated by degree square, as variables that should correlate well with encounter rate, school size and  $f(0)$  (or effective strip width,  $1/f(0)$ ). If these values are used unsmoothed, post-stratification leads to substantial bias, and no estimates are available for degree squares in which there was no effort. It is therefore necessary to use a smoothing routine on each of these variables. This implicitly assumes that each component of Equation 8.1 varies smoothly throughout the stock area. For a given component, the smoothed values for each degree square are ordered from smallest to largest, the number of strata is determined as a function of sample size, and the ordered list divided so that the sample size in each stratum is as equal as possible. The relevant component is then estimated within each stratum, and an average of the stratum estimates, weighted for stratum areas, is calculated. Strata are then determined for the next component, and the process is repeated. Once all component estimates are calculated, they are combined using Equation 8.1. To estimate  $f(0)$  within each stratum, the hazard-rate model was assumed, and data were truncated at five nautical miles.

Criteria were set up and data failing to meet them were discarded. Many complete cruises and substantial parts of many more were deleted in this way, to reduce sensitivity of the methods to bias. Full details are given by Buckland and Anganuzzi (1988) and by Anganuzzi and Buckland (1989). The bootstrap was used to generate robust precision estimates. The bootstrapping unit was taken to be a cruise, to take account of variation between different observers, crews and vessels, while maintaining a fair degree of independence between units. Estimates for 1975–89 for the northern offshore stock of spotted dolphins, taken from Anganuzzi and Buckland (1989) and Anganuzzi *et al.* (1991), are given in Table 8.12.

Buckland *et al.* (1992a) estimated the underlying trends in dolphin abundance by smoothing the estimates of Table 8.12. They considered various smoothing methods such as moving averages, running medians and polynomial regression. Their chosen method was a compound running median known as 4253H, twice (Velleman and Hoaglin 1981).

Suppose that  $\{X(t)\}$ ,  $t = 1, \dots, N$ , is a time series of length  $N$ , and let  $\{S_i(t)\}$  be a smoothed version of it, found by calculating an  $i$ -period running median. We can construct compound smoothing methods such as  $\{S_j(t)\}$ , which is simply  $\{S_j(S_i(t))\}$ . Thus, a 4253 running median method smooths a time series using a 4-period running median, which

### USE OF TUNA VESSEL OBSERVER DATA

**Table 8.12** Estimates of relative abundance and related parameters for the northern offshore stock of spotted dolphin. [ ] indicate that the 1977 estimate was used. Bootstrap standard errors (estimated from  $B = 79$  replicates) are given in parentheses.  $n$  = number of sightings after deletion and truncation;  $\hat{\mu} = 1/f(0)$  (nautical miles);  $n/L$  = estimated encounter rate (schools/1000 n.m.);  $\hat{E}(s)$  = estimated average school size for a random point in the stock area;  $\hat{N}_s$  = estimated number of schools in stock (thousands);  $\hat{N}_d$  = estimated number of dolphins in stock (millions)

Year	$n$	$\hat{\mu}$	$n/L$	$\hat{E}(s)$	$\hat{N}_s$	$\hat{N}_d$
1975	761	[2.74] (0.30)	9.00 (1.28)	634 (71)	6.23 (1.14)	3.95 (1.00)
1976	876	[2.74] (0.30)	7.40 (0.92)	830 (92)	5.13 (0.88)	4.25 (0.91)
1977	1700	2.74 (0.30)	6.46 (0.52)	855 (73)	4.48 (0.69)	3.83 (0.75)
1978	720	2.65 (0.23)	6.58 (0.53)	680 (69)	4.72 (0.55)	3.21 (0.54)
1979	516	2.36 (0.32)	7.03 (0.67)	521 (54)	5.66 (0.92)	2.95 (0.56)
1980	1460	2.37 (0.28)	6.36 (0.39)	654 (80)	5.10 (0.71)	3.34 (0.58)
1981	1593	2.62 (0.24)	6.39 (0.37)	547 (44)	4.63 (0.56)	2.54 (0.44)
1982	1383	2.37 (0.57)	6.32 (0.34)	503 (54)	5.07 (1.34)	2.55 (0.56)
1983	731	2.92 (0.37)	5.94 (0.64)	316 (44)	3.86 (0.53)	1.22 (0.25)
1984	636	3.19 (0.34)	8.81 (0.78)	411 (51)	5.25 (0.74)	2.16 (0.36)
1985	1976	2.78 (0.18)	8.96 (0.63)	471 (41)	6.12 (0.56)	2.88 (0.35)
1986	2197	2.76 (0.17)	9.11 (0.41)	504 (30)	6.28 (0.50)	3.16 (0.30)
1987	3529	2.85 (0.13)	8.82 (0.41)	502 (36)	5.89 (0.42)	2.95 (0.29)
1988	2259	3.15 (0.16)	8.10 (0.62)	550 (35)	4.89 (0.45)	2.69 (0.33)
1989	3569	3.42 (0.18)	8.39 (0.30)	624 (40)	4.66 (0.32)	2.91 (0.28)

is in turn smoothed by a 2-period running median, smoothed again by a 5-period running median, and then by a 3-period running median (i.e.  $\{S_{4253}(t)\} = \{S_3(S_5(S_2(S_4(t))))\}$ ). Near the endpoints, where there are not

## ILLUSTRATIVE EXAMPLES

enough values surrounding a point to be smoothed using the specified running median, a shorter period running median may be used. The endpoints of the resultant time series are calculated by estimating  $X(0)$  and  $X(N + 1)$ , the ‘observed’ values at  $t = 0$  and  $t = N + 1$ , and then calculating

$$S_{4253}(1) = \text{median} \{ \hat{X}(0), X(1), S_{4253}(2) \}$$

and  $S_{4253}(N) = \text{median} \{ S_{4253}(N - 1), X(N), \hat{X}(N + 1) \}$

$\hat{X}(0)$  is found by extrapolating from the straight line which passes through the smoothed values at  $t = 2$  and  $t = 3$ , i.e.  $\hat{X}(0) = 3 \cdot S_{4253}(2) - 2 \cdot S_{4253}(3)$ ; similarly,

$$\hat{X}(N + 1) = 3 \cdot S_{4253}(N - 1) - 2 \cdot S_{4253}(N - 2)$$

The H in 4253H, twice denotes a linear smoothing method commonly used with running medians, which is known as Hanning. It is a 3-period weighted moving average for  $t = 2, \dots, N - 1$ , with weights {0.25, 0.5, 0.25}. The endpoints remain unchanged.

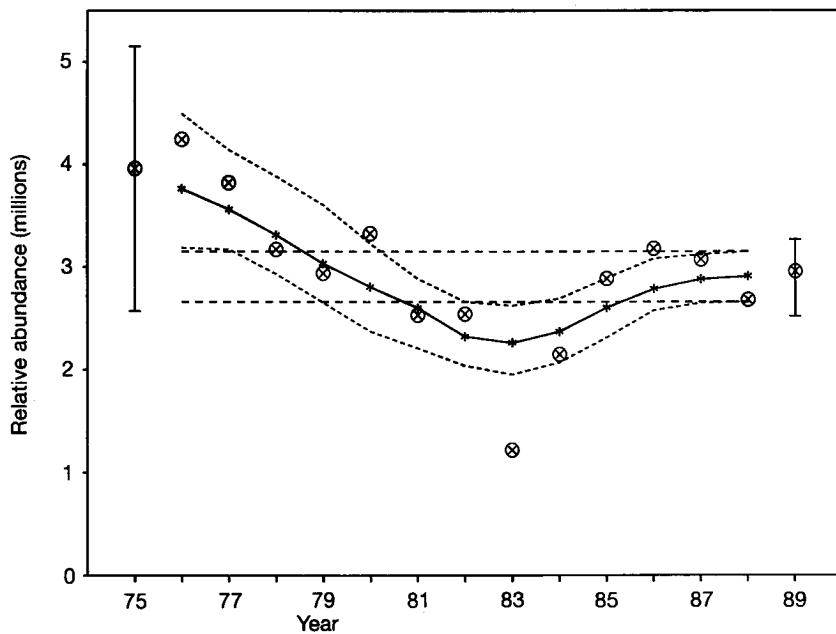
The pattern of the time series may be recovered by calculating the residuals of the series (i.e. the differences between the smoothed and unsmoothed estimates), smoothing the residual series by the same method as the time series was smoothed by, and then adding the smoothed values of the residuals to the smoothed values of the series. This is known as smoothing ‘twice’. For example, if we define the residuals of the time series smoothed by 4253H to be  $\{E(t)\} = \{X(t) - S_{4253}(t)\}$ , then the values of the times series smoothed by 4253H, twice can be defined by

$$\{S_{4253H, \text{twice}}(t)\} = \{S_{4253H}(t) + S_{4253H}(E(t))\}$$

Thus, the 4253H, twice running median method uses a 4253 running median to smooth the time series, estimates the endpoints of the smoothed series and then smooths the resultant series by Hanning. The residuals of the series are calculated and are also smoothed, using the same method as above. The smoothed values of the residuals are then added to the smoothed values of the time series to produce a time series smoothed by 4253H, twice. The advantage of using running medians is that the magnitude of an extreme estimate does not affect the resultant smoothed time series.

To assess changes in abundance over time, the bootstrap was again used, and 85% confidence intervals for relative abundance in each year

## USE OF TUNA VESSEL OBSERVER DATA



**Fig. 8.10.** Smoothed trends in abundance of the northern offshore stock of spotted dolphin. The broken lines indicate approximate 85% confidence limits. The horizontal lines correspond to 85% confidence limits for the 1988 estimate. If these limits both lie above the upper limit for an earlier year, abundance has increased significantly between that year and 1988 ( $p < 0.05$ ); if the limits both lie below the lower limit for an earlier year, abundance has decreased significantly.

were estimated using the percentile method. The rationale for the choice of confidence level is that if two 85% confidence intervals do not overlap, the difference between the corresponding relative abundance estimates is significant at roughly the 5% level ( $p \leq 0.05$ ), whereas if they do, the difference is not significant ( $p > 0.05$ ). One bootstrap replication was carried out for each year, and the bootstrap estimates were smoothed using the running median routine. This process was repeated 79 times, and for each year, the sixth smallest and sixth largest smoothed estimates were taken as approximate 85% confidence limits. The median of the smoothed bootstrap estimates (i.e. the 40th estimate of each ordered set of 79) was used as the 'best' estimate of trend. Figure 8.10 shows the estimates of underlying trend for the northern offshore stock of spotted dolphins. The broken horizontal lines correspond to the upper and lower 85% confidence limits for the 1988 relative abundance

## ILLUSTRATIVE EXAMPLES

estimate. Years for which the entire confidence interval lies outside the region between the broken horizontal lines show a significantly different relative abundance from that for 1988. The estimated trend is downwards until around 1983. Estimated abundance in 1976 was significantly higher than in 1988 ( $p < 0.05$ ), but there is some evidence of a recovery between 1983 and 1988 ( $p \doteq 0.05$ ). Thus northern offshore spotted dolphins appeared to decrease through the 1970s and early 1980s, with numbers remaining stable or increasing since.

### 8.7 House wren densities in South Platte River bottomland

We use data on house wrens (*Troglodytes aedon*) to illustrate model selection. The data were collected from 155 points, with between 14 and 16 points in each of ten 16 ha study blocks. The blocks were established in riparian vegetation along 30 km of South Platte River bottomland near Crook, Colorado. The study was described by Knopf (1986) and Sedgwick and Knopf (1987). The house wren was the most frequently recorded bird, and sample sizes were sufficient to allow estimation by block as well as across blocks. Thus, the option to stratify can also be examined.

The following models were tried: Fourier series (uniform key and up to four cosine adjustments); Hermite polynomial (half-normal key and up to four Hermite polynomial adjustments); half-normal and up to four cosine adjustments; and hazard-rate with at most two simple polynomial adjustments. Terms were tested for inclusion using the likelihood ratio test with a  $p$ -value of 0.05 and DISTANCE option LOOKAHEAD set to two. Intervals for goodness of fit tests were set at 0.0, 7.5, 12.5, 17.5, 22.5, 27.5, 32.5, 42.5, 62.5 and 92.5 m. The largest detection distances were at 90 m. To assess the impact of truncation, the last two intervals were discarded. Thus, the truncation point was 42.5 m, corresponding to 10% truncation of observations. The intervals were chosen to avoid possible favoured rounding distances, such as 10 m or 25 m. We recommend that goodness of fit is not used for model selection, but if it is, we recommend strongly that intervals are set using the option GOF within the ESTIMATE procedure of DISTANCE. The default intervals used by DISTANCE do not take account of rounding to favoured values, and may frequently give spurious significant test statistics.

A summary of results is given in Table 8.13. Note that the log-likelihood and the Akaike Information Criterion (AIC) are of no use for determining whether data should be truncated; values for different models are only comparable if the same truncation point is selected. Density estimates from untruncated data in Table 8.13 are mostly smaller and more precise than those from truncated data. They are

## HOUSE WREN DENSITIES IN SOUTH PLATTE RIVER BOTTOMLAND

probably also more biased given that fits to the untruncated data are less good. The exception is the Hermite polynomial model, which provides the best of the four fits to the untruncated data, and the worst fit to the truncated data. The Fourier series and hazard-rate models perform particularly poorly on the untruncated data. The Fourier series model is not robust to poor choice of truncation point for both line and point transects, whereas the hazard-rate model appears to be robust when data are untruncated for line transects but not for point transects (Buckland 1985, 1987a). The fit of all but the Hermite polynomial model is improved by truncation, and density estimates are more similar under different models for truncated data. We therefore select the model that gives the largest log-likelihood and the smallest AIC value when applied to truncated data for further analyses. This model is the hazard-rate with simple polynomial adjustments.

**Table 8.13** Summary of results from fitting different models to house wren data. FS = Fourier series model (uniform key and cosine adjustments); HP = Hermite polynomial model (half-normal key and Hermite polynomial adjustments); HC = half normal key and cosine adjustments; Hz = hazard-rate key and simple polynomial adjustments. The truncation distance of  $w = 92.5$  m is larger than the largest recorded distance, so no data are truncated, and the value  $w = 42.5$  m corresponds to truncation of 10% of detection distances

Model	Number of adjustments	Log- likelihood	$\chi^2$ statistic (df)	p-value	AIC	$\hat{D}$	Log-based 95% confidence interval
Data untruncated ( $w = 92.5$ m)							
FS	4	- 3312.7	18.8 (3)	< 0.001	6633.4	6.72	(5.95, 7.58)
HP	3	- 3308.4	10.8 (4)	0.03	6624.8	8.28	(6.98, 9.82)
HC	3	- 3308.4	10.7 (4)	0.03	6629.9	8.47	(7.24, 9.91)
Hz	1	- 3329.7	39.3 (4)	< 0.001	6665.5	6.05	(5.28, 6.93)
Data truncated at $w = 42.5$ m							
FS	3	- 2760.0	7.0 (3)	0.07	5526.0	9.05	(7.48, 10.95)
HP	1	- 2762.0	12.1 (4)	0.02	5528.0	7.84	(6.77, 9.07)
HC	1	- 2760.4	7.6 (4)	0.11	5524.8	9.01	(7.43, 10.92)
Hz	1	- 2758.9	7.1 (3)	0.07	5523.8	8.14	(6.44, 10.30)

Estimates stratified by block and by observer are shown in Table 8.14. Goodness of fit tests indicate that fits to eight of the ten blocks are very good, although the data from blocks 0 and 6 are less well modelled. The effective detection radius is high for blocks 0 and 5, but is similar for all other blocks, at around 20 m. Densities vary appreciably between blocks. The final abundance estimate from the analysis stratified by block is very similar to that obtained from an unstratified analysis (Table 8.13, last row). The confidence interval is rather wider, reflecting

### ILLUSTRATIVE EXAMPLES

**Table 8.14** Analyses of house wren data using the hazard-rate model with truncation at 42.5 m. Standard errors are given in parentheses; confidence intervals were calculated assuming a log-normal distribution for  $\hat{D}$ , and the Satterthwaite correction was not applied. \*Estimated density found as the average of the block estimates; the corresponding standard error is found as (square root of the sum of squared standard errors for each block) divided by the number of blocks

Observer	Block	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Estimated density	Log-based 95% confidence interval	Goodness of fit test p-value
All	0	30.4 (1.3)	0.64 (0.14)	2.21 (0.52)	(1.41, 3.47)	0.01
All	1	21.0 (3.2)	1.88 (0.24)	13.60 (4.45)	(7.28, 25.41)	0.42
All	2	22.7 (2.1)	2.17 (0.24)	13.35 (2.87)	(8.81, 20.24)	0.52
All	3	19.1 (10.8)	1.05 (0.14)	9.10 (10.40)	(1.52, 54.46)	0.46
All	4	20.4 (2.2)	1.71 (0.25)	13.05 (3.41)	(7.89, 21.61)	0.35
All	5	31.1 (1.9)	1.20 (0.20)	3.96 (0.83)	(2.64, 5.94)	0.39
All	6	23.5 (2.0)	1.22 (0.15)	7.00 (1.47)	(4.66, 10.51)	0.01
All	7	23.4 (5.1)	1.15 (0.16)	6.70 (3.06)	(2.86, 15.72)	0.35
All	8	23.5 (2.7)	1.16 (0.17)	6.69 (1.80)	(3.98, 11.25)	0.97
All	9	19.3 (4.9)	1.38 (0.22)	11.76 (6.22)	(4.44, 31.15)	0.31
All	All			8.38*	(5.94, 11.83)	
1	All	21.8 (2.6)	1.26 (0.08)	8.44 (2.05)	(5.28, 13.49)	< 0.001
2	All	18.8 (1.8)	1.06 (0.08)	9.56 (1.96)	(6.42, 14.24)	0.84
3	All	19.7 (1.6)	1.12 (0.09)	9.16 (1.67)	(6.43, 13.07)	0.15
4	All	33.0 (1.8)	1.40 (0.10)	4.11 (0.53)	(3.19, 5.30)	0.27
2 and 3	All	19.3 (1.2)	1.09 (0.07)	9.30 (1.28)	(7.10, 12.18)	0.31
1-3	All	20.1 (1.3)	1.15 (0.06)	9.06 (1.22)	(6.96, 11.79)	< 0.001

## HOUSE WREN DENSITIES IN SOUTH PLATTE RIVER BOTTOMLAND

the larger number of parameters that have been estimated. There seems little advantage here to stratification, unless estimates are required by block; this is likely to be true generally when effort per unit area is the same in all strata.

Of more interest is the stratification by observer. Data from observers 1, 2 and 3 yield remarkably similar estimates. However, the first observer's data are modelled poorly. Inspection of the data and output from DISTANCE shows that observer 1 preferentially rounded distances around, or rather over, 10 m to exactly 10 m, and distances between 25 m and 40 m were predominantly recorded as, or close to, 30 m. Such rounding generally generates little bias, but intervals for goodness of fit testing need to be widened and reduced in number to obtain a reliable test when it is present. More serious is the apparent bias in the data of observer 4. The number of detections per point is rather greater than for the other observers, which is consistent with the higher effective detection radius, yet density is estimated to be well under half of that estimated from the data of each of the other observers. It is possible that observer 4 concentrated on detecting birds at greater distances, at the expense of missing many birds close to the point. More likely perhaps is that the effective detection radius was similar to that for the other observers, but that distances were overestimated by observer 4 by roughly 50%. This would be sufficient to explain the large difference in density estimates between observer 4 and the others. Whatever the cause, it seems clear that the data for observer 4 should be viewed with suspicion, whereas those for observers 2 and 3 appear to be most reliable. Our preferred analyses for these data use the hazard-rate model with up to two simple polynomial adjustments and truncation at 42.5 m, are unstratified by block, and discard data from observer 4. If there is concern about the poor fit of this model, the data of observer 1 should also be deleted. The resulting estimates with and without the data of observer 1 are shown in Table 8.14.

We have shown that poor model fits can be improved by truncating data. We now use the untruncated data to illustrate other strategies for improving the fit of a model. Other than truncation, the user of DISTANCE has several options. First, the analyst has control over how many adjustment terms are tested before DISTANCE concludes that no significant improvement in the fit has been obtained. It is not uncommon that a single adjustment term does not improve the fit of a model significantly, whereas the combined effect of it and a further term does yield a significant improvement. If LOOKAHEAD is set equal to one, the better model will not be found, whereas LOOKAHEAD = 2 will allow DISTANCE to select it, at the expense of slower run times. Second, the user can change the method by which DISTANCE selects

### ILLUSTRATIVE EXAMPLES

models using SELECT. For SELECT = **sequential** (the default), it fits the lowest order term first, then adds successively higher order terms sequentially. If SELECT = **forward** is specified, DISTANCE will test for inclusion of each term not yet in the model. If the term that gives the largest increase in the value of the likelihood yields a significant improvement in the fit, it is included, and DISTANCE tests for inclusion of another term. This is analogous to a forward stepwise procedure in multiple regression. SELECT = **all** allows all possible combinations of adjustment terms to be tested, and that giving the minimum Akaike Information Criterion value is selected. Third, the key function (half-normal, hazard-rate, uniform or negative exponential) may be changed, and fourth, a different type of adjustment term (simple or Hermite polynomial, or cosine) can be selected. The combinations of these options that were applied to the house wren data are listed in Table 8.15. The

**Table 8.15** Models and model options used for fitting house wren data. Distance data were pooled across observers and blocks, and were untruncated ( $w = 92.5$  m)

Model	LOOK-AHEAD	Selection method	Key function	Adjustment terms
1	1	Sequential	Hazard-rate	Cosine
2	2	Sequential	Hazard-rate	Cosine
3	1	Forward	Hazard-rate	Cosine
4	-	All	Hazard-rate	Cosine
5	1	Sequential	Hazard-rate	Simple polynomial
6	2	Sequential	Hazard-rate	Simple polynomial
7	1	Forward	Hazard-rate	Simple polynomial
8	-	All	Hazard-rate	Simple polynomial
9	1	Sequential	Uniform	Cosine
10	2	Sequential	Uniform	Cosine
11	1	Forward	Uniform	Cosine
12	-	All	Uniform	Cosine
13	1	Sequential	Uniform	Simple polynomial
14	2	Sequential	Uniform	Simple polynomial
15	1	Forward	Uniform	Simple polynomial
16	-	All	Uniform	Simple polynomial
17	1	Sequential	Half-normal	Cosine
18	2	Sequential	Half-normal	Cosine
19	1	Forward	Half-normal	Cosine
20	-	All	Half-normal	Cosine
21	1	Sequential	Half-normal	Hermite polynomial
22	2	Sequential	Half-normal	Hermite polynomial
23	1	Forward	Half-normal	Hermite polynomial
24	-	All	Half-normal	Hermite polynomial

## HOUSE WREN DENSITIES IN SOUTH PLATTE RIVER BOTTOMLAND

user may also specify the *p*-value for selecting between fits in likelihood ratio tests. For the runs considered in this example, only PVALUE = 0.05 was used, but if this tended to fit too few terms, a larger value (e.g. 0.15) might be preferred.

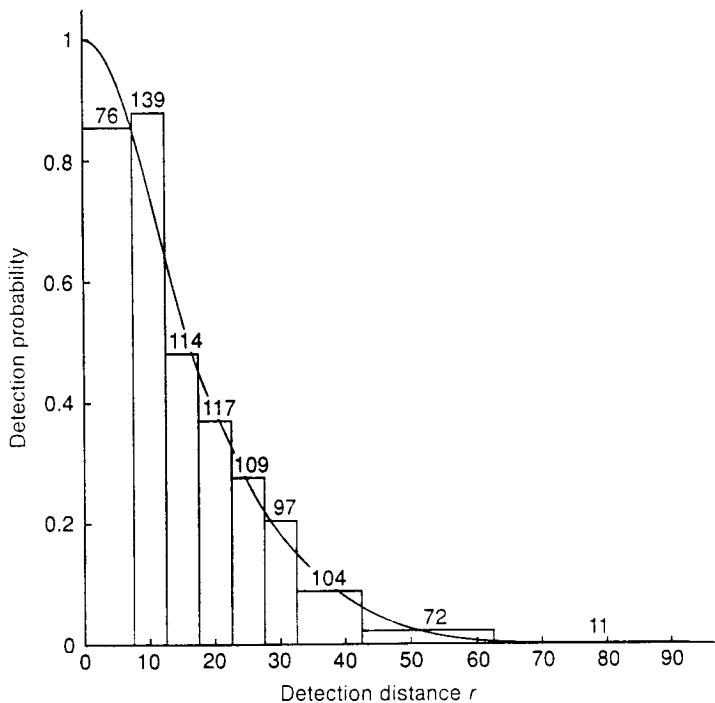
**Table 8.16** Summary of results from fitting different models to house wren data. The models are defined in Table 8.15

Model	Number of adjustments	Log- likelihood	$\chi^2$	(df)	<i>p</i> -value	AIC	$\hat{D}$	Log-based 95% confidence interval
1	0	– 3337.1	61.2	(6)	< 0.001	6678.3	5.56	(4.89, 6.32)
2	3	– 3305.9	8.2	(3)	0.04	6621.9	8.58	(7.33, 10.03)
3 and 4	1	– 3306.7	9.4	(5)	0.10	6619.3	8.74	(7.48, 10.22)
5 and 6	2	– 3321.4	30.1	(4)	< 0.001	6650.8	6.72	(5.79, 7.79)
7 and 8	3	– 3309.8	15.7	(3)	0.001	6629.9	7.51	(6.31, 8.93)
9	3	– 3332.0	59.9	(5)	< 0.001	6670.0	5.28	(4.75, *5.88)
10–12	5	– 3308.1	13.2	(3)	0.004	6626.3	7.46	(6.52, 8.54)
13	3	– 3529.3	501.1	(5)	< 0.001	7064.6	2.65	(2.34, 3.00)
14	5	– 3510.3	306.2	(3)	< 0.001	7030.7	3.24	(2.93, 3.58)
15	3	– 3582.8	515.9	(5)	< 0.001	7171.5	2.56	(2.32, 2.82)
16	4	– 3456.8	271.5	(4)	< 0.001	6921.7	3.39	(3.00, 3.82)
17 and 18	1	– 3313.0	13.8	(6)	0.03	6630.1	7.33	(6.48, 8.30)
19	2	– 3310.4	11.2	(5)	0.05	6626.7	8.38	(7.13, 9.84)
20	3	– 3308.6	10.8	(4)	0.03	6625.2	8.47	(7.24, 9.92)
21 and 23	0	– 3327.3	39.6	(7)	< 0.001	6658.5	6.16	(5.51, 6.89)
22 and 24	2	– 3312.8	13.9	(5)	0.02	6631.7	7.33	(6.39, 8.40)

The results of Table 8.16 indicate a clear ‘winner’ among the models. The hazard-rate model with cosine adjustments and using selection methods forward and all both lead to a hazard-rate model with a single cosine adjustment of order 4. Only this model yields a goodness of fit statistic that is not significant at the 5% level, and its Akaike Information Criterion value is 2.6 lower than the next best model. The density estimate is rather higher than that obtained from the favoured model on truncated data from above. Figure 8.11 shows that the fitted detection function has only a very narrow shoulder. For these data, use of cosine adjustment terms leads generally to a narrow shoulder, whereas polynomial adjustments to the hazard-rate model tend to preserve a wider shoulder. In Fig. 8.12, the fitted detection function obtained by making simple polynomial adjustments to the hazard-rate key, together with selection option forward or all, is shown. Although this model fits the data less well, its wider shoulder may be a better reflection of reality. It yields a density estimate rather lower than that from the favoured method for truncated data from above.

## 8.8 Songbird surveys in Arapaho National Wildlife Refuge

For this example, we use data supplied by F.L. Knopf from extensive songbird surveys of parts of the Arapaho National Wildlife Refuge, Colorado (Knopf *et al.* 1988). We consider counts carried out in June of 1980 and 1981, and analyse the six most numerous species, namely the yellow warbler (*Dendroica petechia*), brown-headed cowbird (*Molothrus ater*), savannah sparrow (*Passerulus sandwichensis*), song sparrow (*Melospiza melodia*), red-winged blackbird (*Agelaius phoeniceus*) and American robin (*Turdus migratorius*). In 1980, three pastures, labelled Pastures 1, 2 and 3, were surveyed by one visit to each of 124, 126 and 123 points respectively. In 1981, four pastures, 0, 1, 2 and 3, were surveyed during one visit to each of 100 points per pasture. All birds detected within 100 m of the point were noted and their locations were flagged, so that their distances could be measured to the nearest 10 cm. Although pastures varied in size, for the purposes of illustration, we assume that each was the same size.



**Fig. 8.11.** The detection function obtained by fitting the hazard-rate key with cosine adjustments to untruncated house wren data.

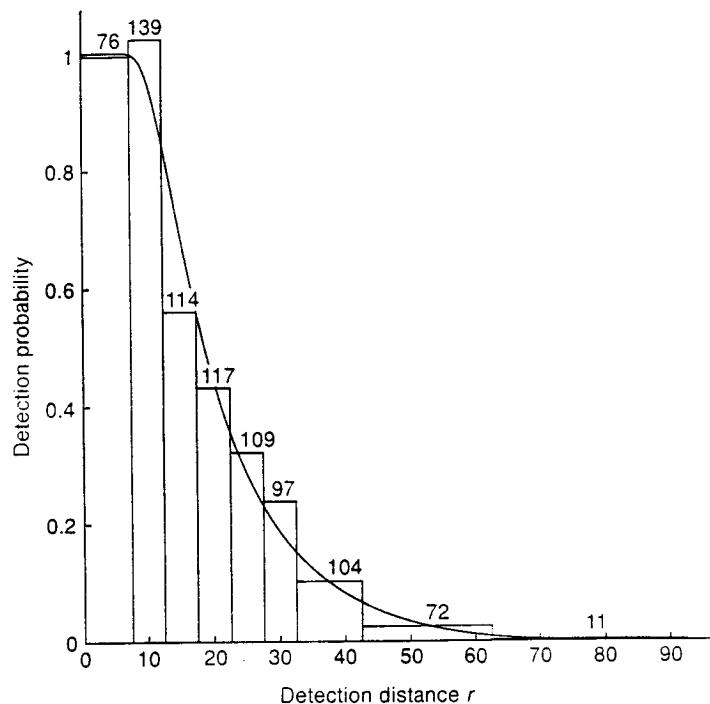


Fig. 8.12. The detection function obtained by fitting the hazard-rate key with simple polynomial adjustments to untruncated house wren data.

Analyses were carried out adopting the half-normal key with cosine adjustments. This model combines the key of the Hermite polynomial model with the adjustments of the Fourier series model. It is computationally more efficient than the former model, and uses a more plausible key function than the latter. For yellow warbler, savannah sparrow and song sparrow, some fits were found to be poor, so the detection distances were truncated at 52.5 m. Other analyses are untruncated. The variance of the number of detections was found using the empirical within DISTANCE.

Yellow warbler analyses are summarized in Table 8.17. Separate estimates were obtained by stratum (pasture). In all, 205 detections were made in 1980 and 342 in 1981. Analyses of the brown-headed cowbird counts (Table 8.18) are less straightforward. First, count frequencies by distance from the point were highly variable in some pastures. Examination of the data showed that this was caused by detections of groups of birds. If more than one bird was recorded at exactly the same distance from the same point, we assume here that the birds comprised a single

### ILLUSTRATIVE EXAMPLES

flock (cluster). In the analyses of Table 8.18, the option OBJECT = cluster was selected. For the first analyses, detection distance data were pooled across pastures, and a single set of estimates per year determined. The second set of analyses are by pasture. The estimates of  $h(0)$  (or equivalently, of the effective radius of detection  $\hat{r}$ ) are imprecise, because the number of detections (as distinct from individual birds) per pasture was low, ranging from 21 (Pasture 1, 1980) to 50 (Pasture 3, 1980). Potentially more serious, bias may be high, as there is little information from which to select an appropriate model when sample size is small. Indeed, the effective radii for 1981 (Table 8.18) indicate appreciably more variability between pastures than can be explained by the standard errors. Either detectability of brown-headed cowbirds varied substantially between pastures or sample size was inadequate at least in some pastures for estimating the effective detection radius with low bias. If the latter explanation is more likely, then pasture estimates with higher precision and lower bias may be obtained by estimating the

**Table 8.17** Analyses of yellow warbler point transect data, Arapaho National Wildlife Refuge. Standard errors are given in parentheses. Estimated density for category 'all' is found as the average of the pasture estimates; the corresponding standard error is found as (square root of the sum of squared standard errors for each pasture) divided by the number of pastures

Year	Pasture	Effective detection radius $\hat{r}$	Encounter rate $n/k$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
1980	1	26.2 (3.5)	0.73 (0.06)	3.40 (0.97)	(1.97, 5.87)	0.67
	2	27.1 (1.7)	0.45 (0.06)	1.96 (0.36)	(1.38, 2.79)	0.42
	3	17.0 (1.6)	0.37 (0.07)	4.02 (1.06)	(2.42, 6.67)	0.80
	All			3.12 (0.49)	(2.30, 4.23)	
1981	0	17.7 (1.4)	0.71 (0.09)	7.18 (1.44)	(4.87, 10.58)	0.60
	1	20.8 (2.4)	0.98 (0.08)	7.22 (1.75)	(4.52, 11.54)	0.07
	2	26.0 (1.5)	0.69 (0.08)	3.25 (0.53)	(2.36, 4.47)	0.13
	3	16.4 (1.2)	0.62 (0.08)	7.34 (1.43)	(5.03, 10.71)	0.19
	All			6.25 (0.68)	(5.05, 7.74)	

## SONGBIRD SURVEYS IN ARAPAHO NATIONAL WILDLIFE REFUGE

effective radius from data pooled across pastures (first section of Table 8.18), and estimating other parameters individually by pasture (second section of Table 8.18). This assumes that detectability does not vary with pasture, and utilizes the fact that average cluster size and expected

**Table 8.18** Analyses of brown-headed cowbird point transect data, Arapaho National Wildlife Refuge. The first set of results was obtained by carrying out an unstratified analysis, the second set by stratifying by pasture, the third set by estimating  $h(0)$  from unstratified distance data and other components separately by pasture, the fourth set by stratifying by cluster size, and the fifth set by correcting for size bias in mean cluster size. Standard errors are given in parentheses

Year	Pasture	Effective detection radius $\hat{r}$	Encounter rate $n/k$	Mean cluster size $\bar{s}$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
<b>Unstratified</b>							
1980	All	38.2 (2.0)	0.26 (0.03)	1.72 (0.11)	0.99 (0.16)	(0.73, 1.35)	0.44
1981	All	34.8 (1.9)	0.35 (0.03)	1.73 (0.11)	1.59 (0.24)	(1.18, 2.13)	0.44
<b>Stratified by pasture</b>							
1980	1	36.6 (5.0)	0.17 (0.04)	1.71 (0.21)	0.69 (0.26)	(0.34, 1.41)	0.46
	2	38.7 (4.2)	0.21 (0.04)	2.04 (0.28)	0.93 (0.29)	(0.50, 1.70)	0.60
	3	38.5 (2.5)	0.41 (0.06)	1.56 (0.12)	1.36 (0.28)	(0.92, 2.02)	0.55
	All				0.99 (0.16)	(0.72, 1.36)	
1981	0	25.5 (3.2)	0.28 (0.05)	1.39 (0.14)	1.91 (0.62)	(1.03, 3.54)	0.23
	1	56.9 (5.1)	0.33 (0.06)	2.03 (0.26)	0.66 (0.19)	(0.38, 1.14)	0.59
	2	42.1 (2.8)	0.42 (0.06)	1.76 (0.18)	1.33 (0.30)	(0.86, 2.06)	0.44
	3	24.9 (3.4)	0.36 (0.05)	1.69 (0.25)	3.12 (1.05)	(1.64, 5.94)	0.66
	All				1.75 (0.32)	(1.23, 2.49)	
<b>Stratified by pasture, except for <math>h(0)</math></b>							
1980	1				0.63 (0.18)	(0.39, 1.13)	
	2				0.95 (0.24)	(0.61, 1.59)	
	3				1.39 (0.26)	(0.97, 2.00)	
	All				0.99 (0.16)	(0.73, 1.35)	

## ILLUSTRATIVE EXAMPLES

**Table 8.18 (Contd.)**

Year	Pasture	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Mean cluster size $\bar{s}$	Estimated density	Log-based 95% confidence interval	Goodness of fit test p-value
1981	0				1.03 (0.25)	(0.67, 1.68)	
	1				1.76 (0.44)	(1.11, 2.90)	
	2				1.95 (0.43)	(1.28, 2.98)	
	3				1.61 (0.38)	(1.07, 2.63)	
	All				1.59 (0.24)	(1.18, 2.13)	
	Stratified by cluster size						
Single birds							
1980	All	36.5 (2.9)	0.16 (0.02)	1.00	0.37 (0.08)	(0.25, 0.56)	0.07
1981	All	32.5 (2.1)	0.22 (0.02)	1.00	0.65 (0.11)	(0.47, 0.90)	0.19
Clusters ( $\geq$ two birds)							
1980	All	40.5 (3.0)	0.11 (0.02)	2.78 (0.15)	0.58 (0.13)	(0.37, 0.89)	0.49
1981	All	39.4 (4.3)	0.13 (0.02)	2.92 (0.20)	0.80 (0.21)	(0.48, 1.32)	0.30
All birds							
1980	All				0.95 (0.15)	(0.70, 1.29)	
1981	All				1.44 (0.23)	(1.05, 1.97)	
Mean cluster size corrected for size bias							
1980	All	38.2 (2.0)	0.26 (0.03)	1.68 (0.10)	0.97 (0.16)	(0.71, 1.33)	0.44
1981	All	34.8 (1.9)	0.35 (0.03)	1.49 (0.07)	1.36 (0.20)	(1.02, 1.81)	0.44

number of detections can be estimated with low bias from small samples, whereas the effective radius often cannot. The approach is described in Section 3.8. Note that care must be taken when estimating variances.

The third section of Table 8.18 shows the estimates obtained from the above approach, assuming the pastures were equal in area. Note how much some of the pasture estimates differ from those found by estimating the effective radius of detection within each pasture. Note also that the estimates for all pastures combined are the same as those for which all data were pooled (first section of Table 8.18). If exactly the same effort (points per unit area) is expended in each stratum, the two methods are equivalent. However, the current method (1) allows separate estimates by stratum, (2) is still valid if effort differs by stratum, and

## SONGBIRD SURVEYS IN ARAPAHO NATIONAL WILDLIFE REFUGE

(3) is preferable to the fully stratified analysis if sample sizes are too small to estimate effective detection radii reliably by stratum, although it assumes that detectability does not vary across strata.

The fourth section of Table 8.18 shows another method of analysing these data. In this case, data were stratified by cluster size (Quinn 1979, 1985). Detections were divided into two categories: single birds and at least two birds. The results suggest that clusters are more detectable than single birds, although the overall estimates of density differ very little from those obtained above.

The final section of Table 8.18 shows adjusted mean cluster size, estimated by regressing logarithm of cluster size on probability of detection, and from this regression, estimating mean cluster size when probability of detection is one. For each cluster, its probability of detection was estimated by substituting its detection distance into the fitted detection function from the analyses of the first section of Table 8.18. The correlation between log cluster size and detection probability was not significant for 1980 ( $r = -0.022$ ,  $df = 96$ ,  $p > 0.1$ ), and estimation was barely affected. For 1981, the correlation was significant ( $r = -0.193$ ,  $df = 137$ ,  $p < 0.05$ ), and estimated cluster size was reduced (1.49 birds per cluster, compared with 1.73 birds per cluster for the detected clusters).

**Table 8.19** Analyses of savannah sparrow point transect data, Arapaho National Wildlife Refuge. Standard errors are given in parentheses

Year	Pasture	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
1980	1	33.4 (2.6)	0.48 (0.07)	1.36 (0.29)	(0.89, 2.06)	0.02
	2	27.1 (1.3)	0.95 (0.10)	4.12 (0.57)	(3.14, 5.40)	0.17
	3	31.2 (2.0)	0.72 (0.07)	2.37 (0.39)	(1.73, 3.26)	0.86
	All			2.62 (0.25)	(2.18, 3.17)	
1981	0	27.0 (2.8)	0.31 (0.05)	1.36 (0.37)	(0.80, 2.29)	0.87
	1	47.0 (6.8)	0.32 (0.06)	0.46 (0.16)	(0.24, 0.89)	0.12
	2	31.6 (2.8)	0.51 (0.08)	1.63 (0.39)	(1.02, 2.58)	0.96
	3	34.6 (3.7)	0.48 (0.07)	1.27 (0.33)	(0.77, 2.11)	0.66
	All			1.18 (0.16)	(0.90, 1.54)	

## ILLUSTRATIVE EXAMPLES

The analyses for savannah and song sparrows presented no special difficulties, and the estimates are given in Tables 8.19 and 8.20, respectively. For red-winged blackbirds (Table 8.21), it was again necessary to analyse the detections as clusters, although average cluster size was small, so bias arising from possible greater detectability of groups of two or more birds would be small. Data were insufficient to stratify, either by pasture or by cluster size.

The final analyses from this example are of American robin (Table 8.22). Again, sample sizes were too small to stratify, but the data presented no additional problems. A single cosine adjustment to the

**Table 8.20** Analyses of song sparrow point transect data, Arapaho National Wildlife Refuge. Standard errors are given in parentheses

Year	Pasture	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
1980	1	27.0 (1.9)	0.38 (0.05)	1.66 (0.33)	(1.12, 2.45)	0.79
	2	23.7 (1.8)	0.40 (0.06)	2.29 (0.49)	(1.51, 3.47)	0.56
	3	23.8 (1.6)	0.41 (0.06)	2.34 (0.46)	(1.59, 3.42)	0.04
	All			2.10 (0.25)	(1.66, 2.65)	
1981	0	31.5 (3.4)	0.43 (0.06)	1.38 (0.36)	(0.83, 2.30)	0.23
	1	32.1 (2.8)	0.47 (0.07)	1.45 (0.33)	(0.94, 2.25)	0.99
	2	24.6 (2.5)	0.39 (0.06)	2.05 (0.51)	(1.26, 3.32)	0.12
	3	23.9 (2.2)	0.29 (0.05)	1.62 (0.42)	(0.98, 2.67)	0.16
	All			1.63 (0.21)	(1.27, 2.08)	

**Table 8.21** Analyses of red-winged blackbird point transect data, Arapaho National Wildlife Refuge. Standard errors are given in parentheses

Year	Pasture	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Mean cluster size $\bar{s}$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
1980	All	27.9 (1.6)	0.18 (0.03)	1.29 (0.09)	0.93 (0.19)	(0.63, 1.37)	0.38
1981	All	32.9 (3.2)	0.17 (0.02)	1.20 (0.07)	0.61 (0.14)	(0.39, 0.97)	0.39

## ASSESSING THE EFFECTS OF HABITAT ON DENSITY

half-normal fit was selected for both the 1980 and the 1981 data. To show the effects of allowing for estimation of the number of adjustments required on the confidence interval for density, both years' data were reanalysed selecting the bootstrap option for estimating the variance of  $\hat{h}(0)$ . The resulting confidence intervals were (0.15, 0.58) and (0.18, 0.64) respectively, wider than the intervals of Table 8.22, as might be expected.

The density estimates of Tables 8.17–8.22 are consistently higher than those of Knopf *et al.* (1988). They used the Fourier series model on squared distances, as recommended by Burnham *et al.* (1980). We no longer recommend this approach, as it can lead to underestimation of density (Buckland 1987a), so the differences between their estimates and those given here might be anticipated.

**Table 8.22** Analyses of American robin point transect data, Arapaho National Wildlife Refuge. Standard errors are given in parentheses

Year	Pasture	Effective detection radius $\hat{p}$	Encounter rate $n/k$	Estimated density	Log-based 95% confidence interval	Goodness of fit test $p$ -value
1980	All	30.2 (2.8)	0.09 (0.02)	0.30 (0.08)	(0.18, 0.50)	0.22
1981	All	36.1 (3.2)	0.14 (0.02)	0.34 (0.08)	(0.22, 0.52)	0.87

## 8.9 Assessing the effects of habitat on density

The design of line and point transect surveys was discussed in Chapter 7. Suppose estimates of object density are required by habitat. The study area should first be stratified by habitat type. Surveys may then be designed within each stratum as described in Chapter 7. A belt of width  $w$ , where  $w$  corresponds to the distance within which say 85–90% of detections are expected to lie, might be defined just within the border of each stratum. If line or point transects are constrained so that they do not lie within the belt, then differences in density between habitats will be easier to detect. If such provision is not made, density will be underestimated in habitat types holding high densities, and overestimated in habitats with low densities. Often, comparisons of density between uniform blocks of habitat and habitat edge are of interest. Point transects are more suited to such comparisons than line transects. As before, habitat edges should be determined by stratifying the area by habitat categories. Points should then be positioned randomly, or systematically (say every 200 m) with a random starting point, along each edge; this may be achieved by envisaging all the edges placed end-to-end and allocating the points along the entire length,

## ILLUSTRATIVE EXAMPLES

before mapping these points back to the actual edges. Detections at these edge points should be recorded according to which side of the edge (i.e. which habitat type) they are in, to allow edge versus centre comparisons within the same habitat type. Thus there would be two analyses of edge points, each taking the fraction of the circle surveyed to be one half. Alternatively, data from each side can be pooled to obtain an estimate of 'edge' density to compare with densities in either or both habitat types, found as described above.

Note that density at a distance  $w$  from the edge may be appreciably different from that at the edge itself. This will not invalidate the above analysis, unless the trend in density is very great, although the estimated detection function will be biased. Provided the assumption  $g(0) = 1$  holds, the method will give a valid estimate of density at the edge. For similar reasons, points from which density away from the edge is estimated could be taken to be at least a distance  $w$  from the edge, rather than say at least  $2w$ , although a larger value might be preferred for other reasons; for example, a value equal to the maximum likely territory diameter could be chosen.

In reality, habitat may be too patchy and heterogeneous to divide a study area into a small number of strata. In this case, density might be better considered as a function of habitat characteristics. One approach to this would be to include habitat information as covariates, as described in Chapter 3, so that the surface representing object density is modelled. We use as an example a different approach. The following is summarized from Bibby and Buckland (1987).

We consider here binomial count data collected during 1983 for a study into bird populations of recently restocked conifer plantations throughout north Wales. In total, 326 points were covered, divided among 62 forestry plots that had been restocked between 1972 and 1981. Further details are given by Bibby *et al.* (1985).

Each detected bird was recorded as to whether it was within or beyond 30 m of the point. The half-normal binomial model of Section 6.2.1 was applied to these data, together with a linear model, for which analytic results are also available, and a single parameter hazard-rate model, with power parameter  $b$  set equal to 3.3, fitted by numerical methods (Buckland 1987a). Table 8.23 (reproduced from Buckland 1987a) shows that the linear model consistently yields higher estimates of densities than does the half-normal model, which in turn yields higher estimates than the hazard-rate model in most cases. Standard errors of these estimates are similar for all three models. Note that all three models give very similar relative densities between species. For example, the ratio of willow warbler density to wren density is estimated as 2.15, 2.13 and 2.19 under the half-normal, linear and hazard-rate models, respect-

## ASSESSING THE EFFECTS OF HABITAT ON DENSITY

ively. Indeed, all three models give exactly the same ranking of species by density. This suggests that binomial counts may be effective for estimating relative density, but yield potentially biased estimates of absolute density. Total counts, which fail to take account of variability in detectability between species or between habitats, give a markedly

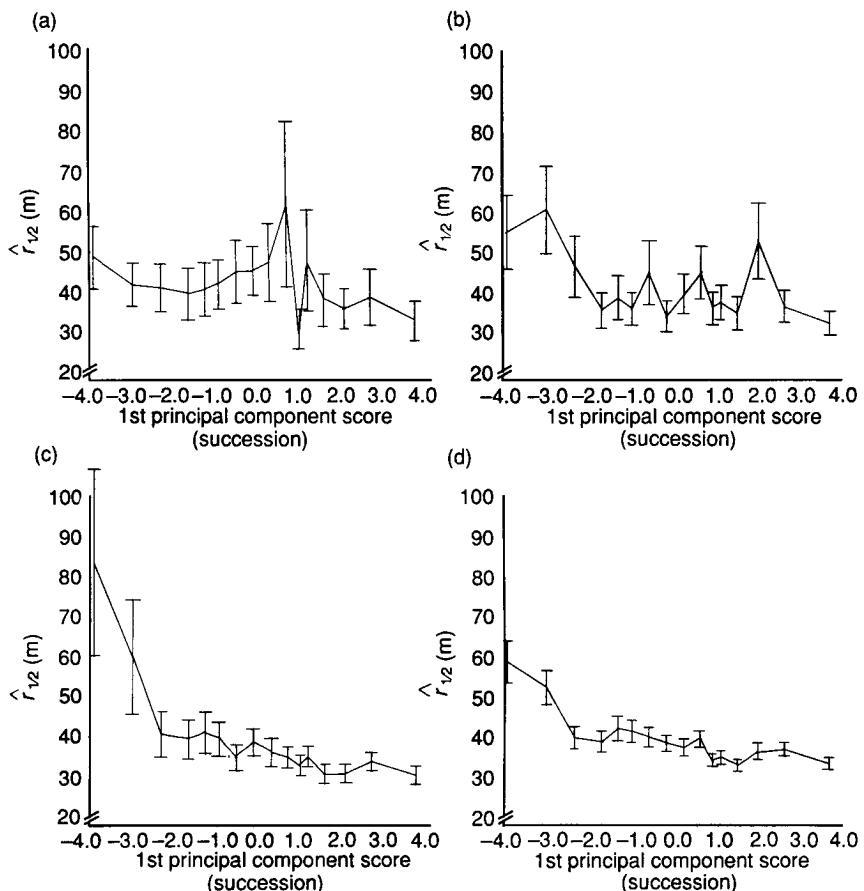
**Table 8.23** Analyses of binomial count data on songbirds from Welsh restocked conifer plantations under three models. Standard errors are given below estimates.  $n_1$  is the number of birds detected within  $c_1 = 30$  m of the point, and  $n_2$  is the number beyond 30 m. Scientific names are given in Appendix A

Species	$n_1$	$n_2$	Linear model			Half-normal model			Hazard-rate model		
			$\hat{D}$	$\hat{r}_{1/2}$	$\hat{p}$	$\hat{D}$	$\hat{r}_{1/2}$	$\hat{p}$	$\hat{D}$	$\hat{r}_{1/2}$	$\hat{p}$
Willow warbler	421	504	6.09 0.37	32.1 0.8	38.5 1.0	6.65 0.38	31.9 0.7	36.8 0.9	5.23 0.34	32.2 1.0	43.3 1.3
Wren	208	347	2.83 0.22	36.4 1.3	43.8 1.5	3.12 0.23	36.1 1.2	41.7 1.4	2.39 0.18	37.3 1.4	50.1 1.9
Goldcrest	108	57	1.90 0.26	24.2 1.2	29.1 1.5	1.96 0.25	24.8 1.1	28.6 1.2	1.99 0.38	21.7 1.9	29.1 2.5
Tree pipit	127	235	1.70 0.18	38.0 1.7	45.6 2.0	1.88 0.19	37.6 1.6	43.3 1.9	1.44 0.15	39.0 1.8	52.3 2.5
Robin	78	89	1.14 0.16	31.5 1.8	37.8 2.2	1.24 0.16	31.4 1.7	36.2 2.0	0.98 0.15	31.5 2.2	42.4 3.0
Chaffinch	73	141	0.97 0.13	38.7 2.3	46.4 2.7	1.07 0.15	38.2 2.2	44.1 2.5	0.82 0.11	39.7 2.4	53.3 3.3
Garden warbler	58	87	0.80 0.13	34.9 2.3	42.0 2.8	0.88 0.13	34.6 2.2	40.0 2.6	0.68 0.11	35.6 2.6	47.8 3.5
Siskin	36	74	0.47 0.10	39.7 3.3	47.6 4.0	0.52 0.11	39.2 3.2	45.3 3.7	0.40 0.08	40.7 3.5	54.7 4.7
Whitethroat	33	48	0.46 0.10	34.5 3.0	41.5 3.7	0.51 0.10	34.2 2.9	39.5 3.3	0.39 0.08	35.1 3.5	47.2 4.6
Coal tit	29	38	0.41 0.10	33.2 3.1	39.8 3.7	0.45 0.10	33.0 2.9	38.1 3.4	0.35 0.09	33.6 3.7	45.1 4.9
Dunnock	28	32	0.41 0.10	31.5 3.0	37.8 3.6	0.45 0.10	31.4 2.8	36.3 3.3	0.35 0.09	31.6 3.7	42.4 4.9
Song thrush	27	79	0.34 0.08	46.1 4.4	55.3 5.3	0.38 0.08	45.5 4.4	52.5 5.1	0.30 0.06	47.4 4.6	63.6 6.1
Long-tailed tit	18	12	0.30 0.10	26.1 3.2	31.3 3.8	0.31 0.10	26.5 2.8	30.5 3.3	0.29 0.12	24.4 4.6	32.8 6.2
Blackbird	15	40	0.19 0.06	44.3 5.7	53.2 6.9	0.21 0.06	43.7 5.6	50.4 6.5	0.16 0.04	45.5 5.9	61.1 7.9
Blackcap	10	6	0.17 0.07	25.2 4.1	30.3 5.0	0.18 0.07	25.7 3.7	29.7 4.2	0.17 0.10	23.2 6.2	31.1 8.3
Redpoll	12	20	0.16 0.07	36.4 5.3	43.8 6.4	0.18 0.08	36.1 5.1	41.6 5.9	0.14 0.06	37.3 5.8	50.0 7.8
Chiffchaff	9	27	0.11 0.04	46.6 7.8	55.9 9.4	0.12 0.05	46.0 7.7	53.1 8.9	0.10 0.03	47.9 8.0	64.3 10.8
Mistle thrush	6	41	0.07 0.03	67.6 13.8	81.2 16.6	0.08 0.04	67.1 14.0	77.5 16.2	0.06 0.03	70.9 16.5	95.2 22.1

## ILLUSTRATIVE EXAMPLES

different ordering of species. Counts of birds within 30 m of the point give a better indication of relative densities, although the ordering of goldcrest (*Regulus regulus*) and tree pipit (*Anthus trivialis*), and of blackcap (*Sylvia atricapilla*) and redpoll (*Carduelis flammea*), is reversed relative to the density estimates.

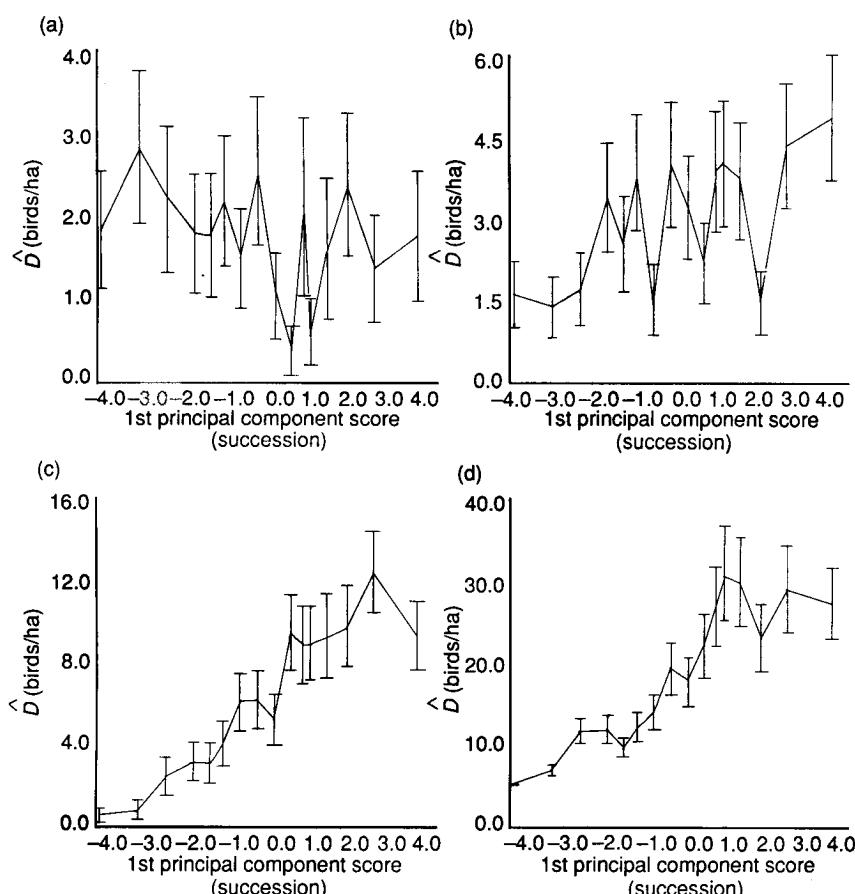
Table 8.23 shows that estimates of  $r_{1/2}$  agree remarkably well under the three models, considering their widely differing shapes, indicating that the estimates provide a useful guide to the relative detectability of species. Variation within a species in values of  $\hat{\rho}$  suggests that estimation of  $\rho$  may be less robust than that of  $r_{1/2}$ .



**Fig. 8.13.** Variation in detectability of (a) tree pipit, (b) wren, (c) willow warbler and (d) all species (pooled data) with habitat succession in conifer plantations aged between 2 and 11 years.

## ASSESSING THE EFFECTS OF HABITAT ON DENSITY

Various aspects of the habitat within a 30 m radius of each point were recorded, and a principal components analysis carried out. The first component was identified as succession. The plantations had been restocked between two and eleven years previously, so the environment ranged from open to very dense. Birds of each species were recorded according to whether they were within or beyond 30 m of the point. The binomial half-normal model for the detection function was assumed, and  $r_{1/2}$ , the distance at which probability of detection is one half, was used as a measure of detectability. Three species, the tree pipit, wren (*Troglodytes troglodytes*) and willow warbler (*Phylloscopus trochilus*)



**Fig. 8.14.** Variation in density of (a) tree pipit, (b) wren, (c) willow warbler and (d) all species (pooled data) with habitat succession in conifer plantations aged between 2 and 11 years.

## ILLUSTRATIVE EXAMPLES

were present in sufficient numbers at each stage of development to examine their change in detectability and density with succession in habitat. Figure 8.13 shows the estimated change in detectability with succession for these three species, and for all species combined. Both the wren and the willow warbler appear to be more detectable in the very early stages of succession. The pattern for the tree pipit is less clear. Analyses of the combined data set show a similar pattern to those for wren and willow warbler.

To measure trends in bird density with succession, it is therefore necessary to adjust for greater detectability in more open habitats. In Fig. 8.14 estimated change in density with succession is shown for the same three species and for all species combined. Both the wren and the willow warbler show a trend to higher densities in the older plantations. The plot for all species combined shows roughly a fivefold increase in density for eleven year old restocks relative to two year old. If unadjusted counts of birds are used as measures of relative abundance, this increase is estimated to be just 1.4-fold, indicating the importance of adjusting counts for detectability.

Principal components analysis was used in the above because it proved effective at reducing the dimensionality of the habitat variables. The second component represented a trend from a more diverse habitat, with herbaceous plants and regenerating broadleaf trees, through to pure coniferous stand with little undergrowth. If the only aspects of interest were variation in detectability and density with succession, the analysis could have been simplified by replacing the first principal component by stand age.