NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

International Commission for the Northwest Atlantic Fisheries

Serial No. 5474

ICNAF Res. Doc. 79/VI/109

ANNUAL MEETING - JUNE 1979

TRANSECT SAMPLING METHODS AND THEIR APPLICATIONS TO THE DEEP-SEA RED CRAB

by

G. P. Patil and C. Taillie
Department of Statistics, The Pennsylvania State University, and
The Satellite Program in Statistical Ecology, International Statistical Ecology Program
State College, PA 16802

and

R. L. Wigley
National Marine Fisheries Service
Northeast Fisheries Center
Woods Hole, Massachusetts 02543

ABSTRACT

Line transect techniques were employed in a survey to determine the density of the deep-sea red crab, *Geryon quinquedens*. Photographs of the sea bottom were taken in continental slope waters off northeastern U. S. from offshore Maryland (38°N., 74°W.) to Georges Bank (41.5°N., 66°W.) in July 1974. Water depths at the 33 sampling stations ranged from 210 to 1,463 meters. Photographs were taken with a sled-mounted underwater camera system towed along the sea bottom by R/V *Albatross IV*. Some of the basic mathematical concepts on which this survey was based are analyzed and show the utility of the line transect method.

PART A: INTRODUCTION

A line transect is a survey technique wherein the number of organisms of a given species are recorded as observed along a path, usually a straight line, through the survey area. The National Marine Fisheries Service applied this technique in conducting a quantitative photographic survey of the deep-sea red crab, *Geryon quinquedens* Smith, off northeastern United States in June-July 1974. Survey operations were conducted in continental slope waters at depths between 229 and 1,646 m in the area extending from offshore Maryland (38°N., 74°W.)

northeastward to the eastern end of Georges Bank (41.5°N., 66°W). The primary purpose of the survey was to obtain quantitative estimates of the number and biomass of red crabs in that region. Secondary purposes were to assess the size composition of this species, and to obtain additional information on its distribution, life history, and ecology. General information resulting from this survey was reported by Wigley, Theroux, and Murray (1975). Emphasis in that report was placed on the quantitative distribution of the red crab in terms of density and biomass; ecological and statistical aspects were secondary. The purpose of the present report is to review the statistical aspects of line transect theory as it pertains to this survey.

Materials and Methods

The survey was conducted from aboard the research vessel ALBATROSS IV, a 57-m vessel operated for the Northeast Fisheries Center by the National Ocean Survey, Office of Fleet Operations, NOAA. Sampling gear of two types was used on this survey: (1) an underwater photographic system (Figure 1), which took *in situ* photographs of the sea bottom and constituent epibenthic fauna, and (2) an otter trawl used to catch red crabs.

Photographs of the sea bottom were obtained for the purpose of determining the density of red crabs. The photographic system consisted of a 70-mm camera and stroboscopic light mounted on a large steel sled. Dimensions of the sled are: 2.7 m long, 2.1 m wide, and 1.9 m high. It was constructed of heavy-gauge 6.4-cm diameter steel pipe, and runners 25.4 cm broad and 2.5 cm thick. The sled weighs 1,225 kg. The camera was a Hydro-Products Deep Sea Photographic Camera, Model PC-705; the strobe unit was a Hydro-Products, Deep Sea Strob, Model PF-730. Film used was Kodak Ektachrome EF daylight (color) and Kodak Tri-X Pan (black and white).

Stations were pre-selected according to a stratified random design. Previous information indicated that crab densities could be expected to be highest at water depths between approximately 250 and 500 m. Thus, a higher proportion of sampling stations were scheduled for that bathymetric zone. At each station where the sea floor was suitable the photo sled was towed at a speed of 1 to 2 knots. Suitability of the bottom was evaluated by means of echo-sounding just

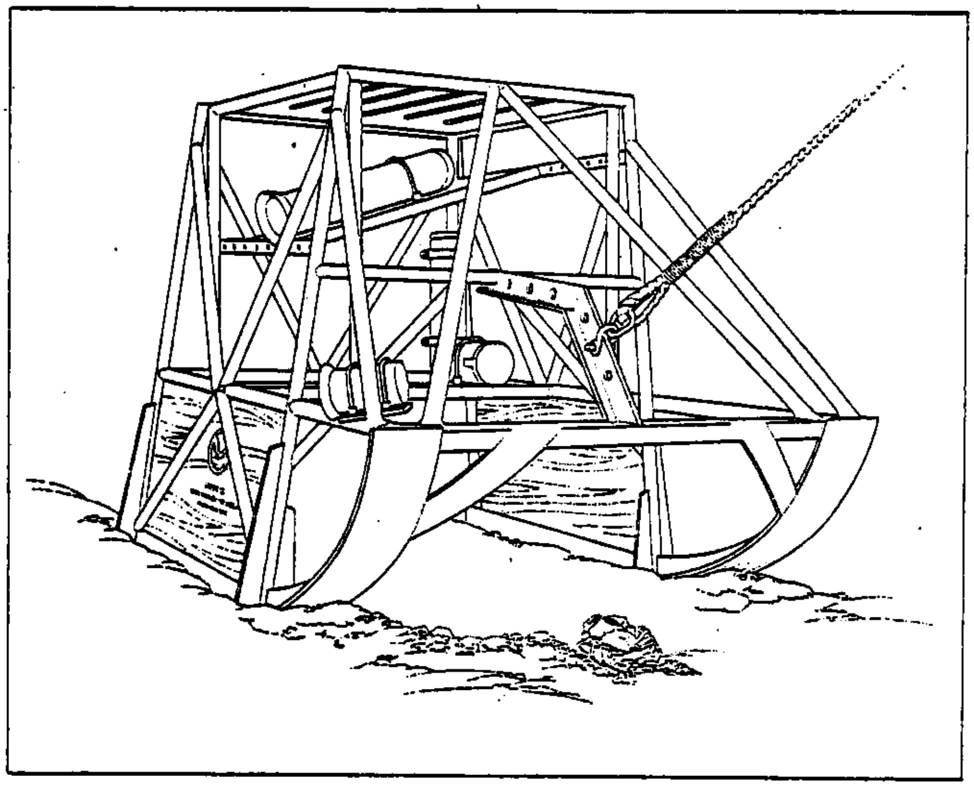


Figure 1. Sled-mounted photographic system. Camera is mounted at the upper right-center of sled (reader's left), the electronic flash unit is in the forward corner, the power pack is fastened horizontally at the left center of the sled, and the orientation pinger is mounted vertically in the rear corner.

prior to launching the photo sled. At shallower depths, less than 585 m, an ELAC fathometer was used and in deeper water an EDO instrument was used. The tow duration of the photo sled at each station ranged from 30 to 75 minutes, depending on local conditions (bottom roughness, anchored fishing gear, etc.). The camera was programmed to obtain a photograph every 10 seconds; thus the maximum number of photographs during one tow was approximately 400. Upon completion of the photo-sled tow, the film was removed, and a short strip (1 to 2 m) was developed to monitor focus, strobe light position, exposure, etc. The remainder of the film was brought to the laboratory ashore and sent to commercial film processors for developing and printing. A total of 18,000 photographs was obtained at 33 stations. Of this total, 8,262 photographs representing the best quality for enumeration purposes were selected for quantitative analyses.

PART B: STATISTICAL METHODOLOGY

In the usual line transect setup, the observer moves along a transect and records the number of organisms sighted. However, not all organisms present in the region under study may be actually observed. In fact, those further away from the transect are less likely to be sighted. Thus observed density tends to be an underestimate of the true density of the organisms. In order to remove this visibility bias, the investigator may either inflate the recorded count of the organisms or replace the area of the region by something smaller, called effective area. The line transect theory provides a method to determine the effective area by using right angle distances to the organisms sighted, for example, see Burnham and Anderson (1976), Eberhardt (1978), Gates (1968, 1978), and Seber (1973). This is achieved by introducing the concept of visibility function $g(x)$ defined as the probability of sighting an organism present at right angle distance x. The shape of the visibility function $g(x)$ is usually reflected in the empirical graph of the number of sighted organisms against their right angle distances. In accord with intuition, these graphs are usually decreasing.

For the data under study in this paper, the counts of organisms first increase and then decrease with the right angle distances. This unusual feature can be, however, explained by decreasing visibility coupled with progressivity expanding field of view of the camera. This composite effect can be represented by what we call a weighted visibility function. In what follows, we use this concept to develop certain methods for estimating population density.

Weighted Visibility Function and Effective Area

The camera was mounted on the sled at an angle with the horizontal. As a consequence of this, what appears to be a rectangle in the photograph is in fact a trapezoid on the ocean floor. This is displayed in Figure 2. Let a and H be the base and the height of the trapezoid. The length of the trapezoid at distance x is of the form $a + bx$, where b is a dimensionless constant. For our setup, the numerical values of a, b, and H are 2.868m, .961, and 5.49m respectively.

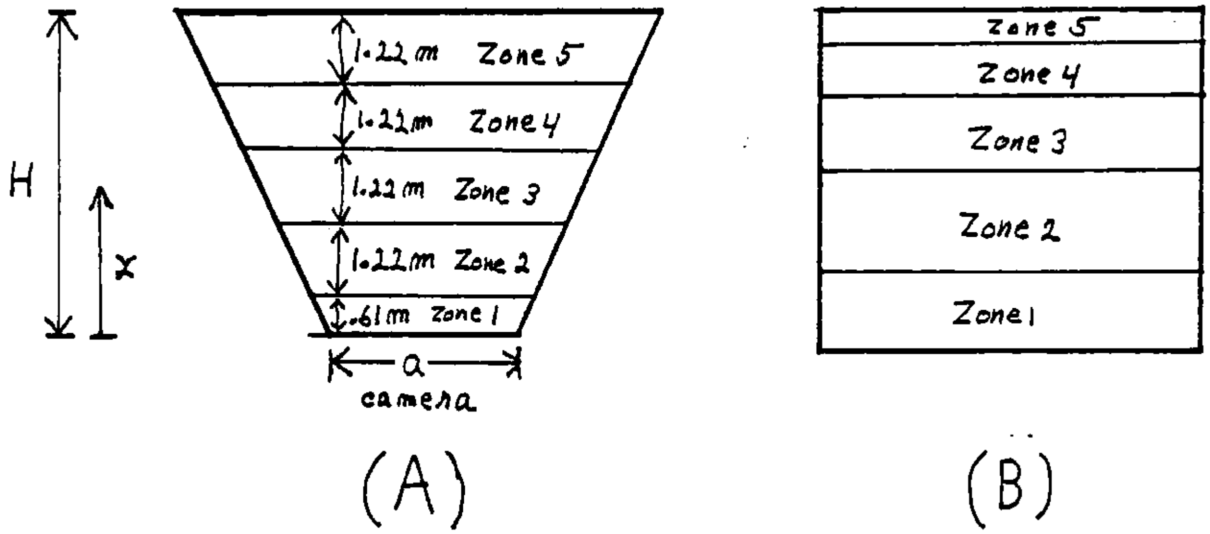


Figure 2: Schematic diagrams of (A) the actual field of view on the ocean floor and (B) its photographic image.

When a crab was detected on the photograph, its right angle distance x was measured by superimposing a grid and counting the number of squares to the crab. This procedure was, however, found to be tedious and time-consuming. As an alternative to measuring the distances, the field of view was divided into five zones as indicated in Figure 2, and the number of crabs was counted for each zone. Originally all the five zones were of equal width. But it was found that, near the sled, sediment clouds obscured the visibility, and therefore, only the upper half of the first zone has been analyzed.

Now, if $g(x)$ is the visibility function, and $n(x)$ is the number of crabs sighted in the trapezoid strip $(x, x+\Delta x)$,

$$E[n(x)] = s.D.\ (a+bx)\Delta x \cdot g(x) , \qquad (1)$$

where s is the number of frames (photographs) and D is the density. This implies that the probability density function $f(x)$ of recorded x is proportional to $(a+bx)g(x)$ giving

$$f(x) = \frac{(a+bx)g(x)}{w} \qquad 0 \leq x \leq H \qquad (2)$$

where $$w = \int_0^H (a+bx)g(x)dx. \qquad (3)$$

Further, from equation (1) we can estimate density D by

$$\frac{n(x)}{s(a+bx)g(x)\Delta x} .$$

As is well known, the minimum variance pooled estimate is then given by

$$\hat{D} = \frac{n}{s \int_0^H (a+bx)g(x)dx} = \frac{n}{sw} , \tag{4}$$

where n is the total number of crabs recorded on s frames. Because of the nature of equation (4), w is called the effective area per frame.

In order to utilize (4), we need to know w , which appears as a parameter in the probability density function of the recorded right angle distances as given in equation (2). The parameter w can be estimated in principle from these distances. We discuss two possible methods in the next two sections.

Exponential Estimators

The exponential sighting function has received considerable attention in the line transect literature. For our purposes in the present study we use

$$g(x) = e^{-\lambda x} , \qquad 0 \leq x \leq H . \tag{5}$$

The graph of the corresponding weighted visibility function $(a+bx)e^{-\lambda x}$ increases for small x and decreases for large x as needed in our present study. While the appropriateness of the exponential sighting function for our data is discussed later, we provide the necessary results for estimating the density here.

The effective area is given by

$$w = \frac{a}{\lambda} (1 - e^{-\lambda H}) + \frac{b}{\lambda^2} (1 - e^{-\lambda H} - \lambda H e^{-\lambda H}) . \tag{6}$$

The maximum likelihood estimate of λ is the solution of

$$\bar{x} = - \frac{1}{w} (\frac{dw}{d\lambda}) = \frac{2}{\lambda} - \frac{1}{w} [\frac{a}{\lambda^2} (1 - e^{-\lambda H}) + \frac{(a+bH)H}{\lambda} e^{-\lambda H}] \tag{7}$$

where $\bar{x}$ is the average x.

Using the method of statistical differentials (see Johnson and Kotz, 1969, p.29), the asymptotic variance of $\hat{D}$ in (4) is given by

$$\mathrm{Var}(\hat{D}) = \frac{D}{sw}\left[\frac{\mathrm{Var}(n)}{E[n]} + \frac{\left(\frac{dw}{d\lambda}\right)^2}{\frac{d^2w}{d\lambda^2}w - \left(\frac{dw}{d\lambda}\right)^2}\right] \quad . \tag{8}$$

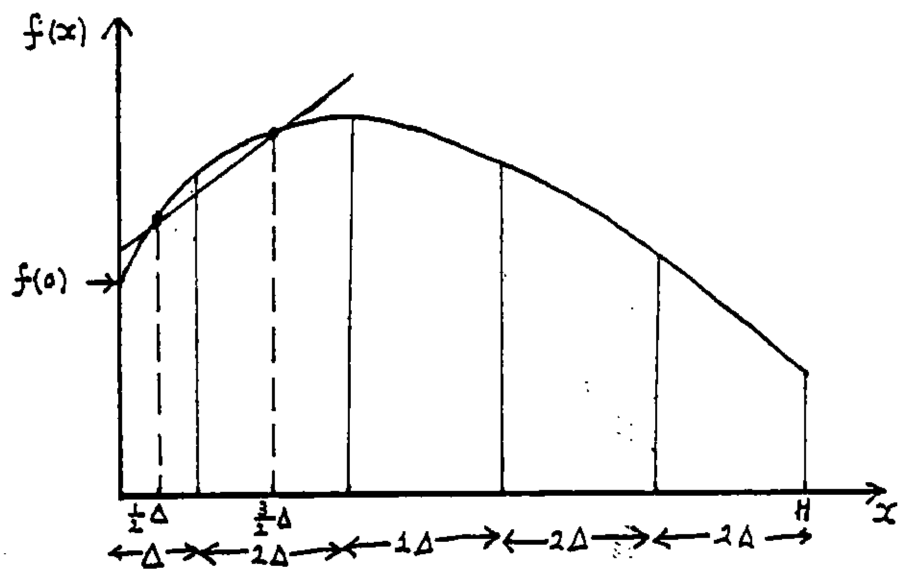
The variance to mean ratio is known to be related to spatial pattern. It is one for Poisson pattern and exceeds one for aggregated pattern. Although there is indication of some aggregation in our data, in the analysis section we have set the variance to mean ratio to be one, thus giving an underestimate for the asymptotic variance of $\hat{D}$. We feel, however, that the bias is not serious.

Cox Estimators

As Eberhardt (1978) has discussed, the Cox estimators provide estimates of the density when no parametric form is assumed for the sighting function. Putting $x = 0$ in equation (2) gives

$$f(0) = \frac{a}{w} \tag{9}$$

provided $g(0)$ can be assumed to be unity. The Cox method obtains the estimate of w by estimating $f(0)$ as follows.



Consider the first two zones with widths Δ and 2Δ respectively. Let n_1 and n_2 be the observed counts of the crabs in these zones. Then $f(\frac{1}{2}\Delta)$ and $f(\frac{3}{2}\Delta)$ are estimated by $\frac{n_1}{n\Delta}$ and $\frac{n_2}{2n\Delta}$. The linear extrapolation back to the origin gives the estimate

$$\hat{f}(0) = \frac{8n_1 - n_2}{6n\Delta} \; . \tag{10}$$

Combining (4), (9) and (10) gives Cox estimator

$$\hat{D} = \frac{1}{sa\Delta} \left(\frac{8n_1 - n_2}{6}\right) \; . \tag{11}$$

Under random spatial pattern, n_1 and n_2 are independent Poisson. It follows that $Var(\hat{D})$ can be estimated by

$$\hat{Var}(\hat{D}) = \frac{1}{(sa\Delta)^2} \left(\frac{64\ n_1 + n_2}{36}\right) \tag{12}$$

PART C: DATA ANALYSIS

<u>Density Estimation by Zone Counts</u>

We illustrate the methodology of Part B using zone counts from three stations (16, 21, 67). Computations and estimates are summarized in Table 1. Note that $\bar{x}$ is computed from grouped data in the form of zone counts. The effect of grouping on the estimates is not sizable, as discussed in a later section.

We have assumed the exponential visibility function for parametric approach. This assumption appears reasonable, at least according to the chi-square test. The Cox estimates are based on non-parametric procedures that do not assume a parametric form for the visibility function. It is interesting to compare the exponential estimates with Cox estimates. As expected of non-parametric procedures, the Cox estimates have larger standard errors. Further we observe that the two density estimates are close for Station 16. For Station 21, the Cox estimate is smaller than the exponential estimate, whereas the reverse is true for Station 67. For this reason, it may be worthwhile to examine more closely the exponentiality of the visibility function.

TABLE 1: DENSITY ESTIMATION USING ZONE COUNTS AND EXPONENTIAL SIGHTING FUNCTION. (ESTIMATES ARE ± ONE STANDARD DEVIATION)

Station	Photo Zone	Zone Mid-point (Meters)	Number of Crabs Observed	Expected	Estimates
16	1	.305	11	10.7	Exponential Estimates
Depth: 530–530 m	2	1.22	22	22.5	$\hat{\lambda}$ = .207 m^{-1}; $\hat{w}$ = 16.47 m^2
	3	2.44	22	22.6	$\hat{D}$ = 149 ± 30 crabs/ha
394 frames	4	3.66	23	21.5	Cox Estimates
	5	4.88	19	19.8	$\hat{w}$ = 15.42 m^2; $\hat{D}$ = 160 ± 65 crabs/ha
		$\bar{x}$ = 2.69	97	χ^2_3 = .17	
21	1	.305	17	20.5	Exponential Estimates
Depth: 393–412 m	2	1.22	40	35.9	$\hat{\lambda}$ = .414 m^{-1}; $\hat{w}$ = 9.92 m^2
	3	2.44	34	28.0	$\hat{D}$ = 299 ± 48 crabs/ha
404 frames	4	3.66	12	20.7	Cox Estimates
	5	4.88	17	14.8	$\hat{w}$ = 13.12 m^2; $\hat{D}$ = 226 ± 79 crabs/ha
		$\bar{x}$ = 2.20	120	χ^2_3 = 6.3	
67	1	.305	18	14.0	Exponential Estimates
Depth: 412–960 m	2	1.22	26	26.5	$\hat{\lambda}$ = .323 m^{-1}; $\hat{w}$ = 12.24 m^2
	3	2.44	18	23.1	$\hat{D}$ = 189 ± 36 crabs/ha
423 frames	4	3.66	17	19.1	Cox Estimates
	5	4.88	19	15.3	$\hat{w}$ = 8.72 m^2; $\hat{D}$ = 266 ± 77 crabs/ha
		$\bar{x}$ = 2.41	98	χ^2_3 = 3.5	

Appropriateness of Exponential Visibility Function

For each zone, the observed and expected number of crabs were divided by the zonal area. These densities are plotted in Figure 3 against the zone mid-points. The observed and expected points closely match for Station 16. For Station 21, the observed plot appears to be more like half-normal than exponential.

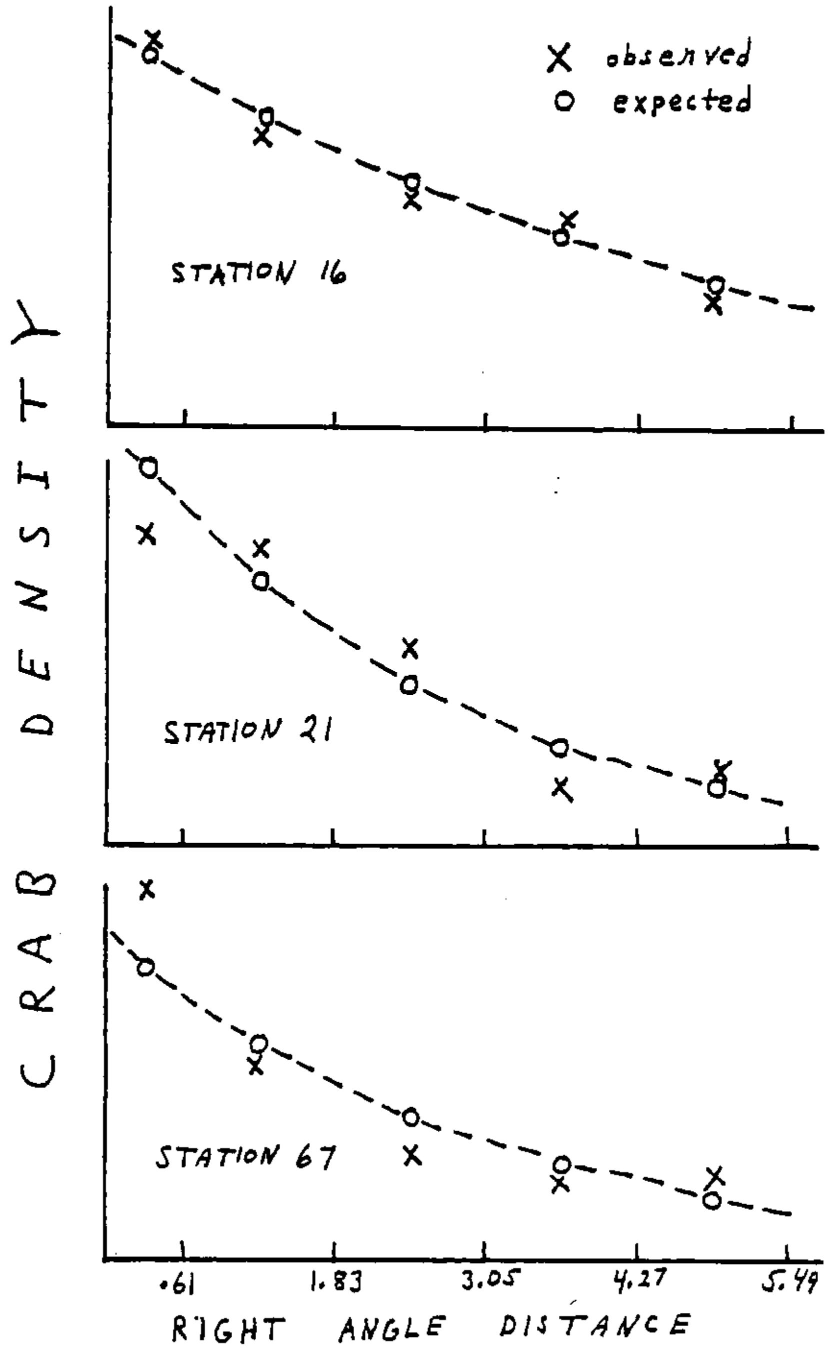


Figure 3: Observed and expected crab density versus right angle distance at three stations.

The use of exponential sighting function in such a case tends to over-estimate the population density. This may partly explain the smaller value of the non-parametric Cox estimate for Station 21. Examination of the plots for

Station 67 reveals that observed points fall off more rapidly and are more convex than the expected points. In this case, the exponential estimate is likely to be an underestimate. This may partly explain the larger value of the non-parametric Cox estimate for Station 67.

In view of the above observations, it would be more desirable to work with a richer and more flexible family of visibility functions which includes a wide variety of shapes. A promising possibility lies in the exponential-power-family defined by

$$g(x) = \exp(-\lambda x^{\gamma})$$

This family includes the exponential ($\gamma=1$), half-normal shapes ($\gamma>1$), and shapes more convex than the exponential ($0<\gamma<1$). Estimation procedures for this family are presently under investigation.

Effect of Grouping

For the three stations under consideration, some data was also available in the form of measurements of right angle distances. This has enabled us to study the effect of grouping referred to earlier. This study is summarized in Table 2. It shows that grouping tends to underestimate the density estimates and their standard errors. The bias introduced in the density estimates seems, however, small for the sizes of the standard errors. Thus, one may choose to work with zonal count data and not expend any effort in measuring right angle distances. In this context, the question of 'how many zones' should be examined further.

References

Burnham, K. P. and Anderson, D. R. (1976). Mathematical models for non-parametric inferences from line transect data. Biometrics 32: 325-336.

Eberhardt, L. L. (1978). Transect methods for population studies. J. Wildlife Management 42: 1-31.

Gates, C. E. (1968). Line transect method of estimating grouse population densities. Biometrics 24: 135-145.

Gates, C. E. (1978). Line transects and related issues. In Sampling Biological Populations. (Ed. R. Cormack, G. P. Patil, and D. S. Robson), Satellite Program in Statistical Ecology, International Co-operative Publishing House, Fairland, Maryland. pp. 1-100. To appear.

Johnson, N. L. and Kotz, S. (1969). Discrete Distributions. Wiley, New York.

Seber, G. A. F. (1973). The Estimation of Animal Abundance. Hafner Press, New York.

Wigley, Roland L., Roger B. Theroux, and Harriett E. Murray. (1975) Deep-Sea Red Crab, Geryon quinquedens, Survey off Northeastern United States. Marine Fisheries Review, Vol. 37., No. 8, pp. 1-21, August 1975.

TABLE 2: GROUPING EFFECT ON EXPONENTIAL ESTIMATES

Station 16: 51 Crabs

Ungrouped Estimates	Grouped Estimates
$\bar{x}$ = 2.23 m	$\bar{x}$ = 2.26 m
$\hat{\lambda}$ = .400 m^{-1}	$\hat{\lambda}$ = .387 m^{-1}
$\hat{w}$ = 10.24 m^2	$\hat{w}$ = 10.55 m^2
$\hat{D}$ = 240 ± 55 crabs/ha	$\hat{D}$ = 233 ± 54 crabs/ha

Station 21: 93 Crabs

Ungrouped Estimates	Grouped Estimates
$\bar{x}$ = 2.21 m	$\bar{x}$ = 2.30 m
$\hat{\lambda}$ = .410 m^{-1}	$\hat{\lambda}$ = .370 m^{-1}
$\hat{w}$ = 10.01 m^2	$\hat{w}$ = 10.96 m^2
$\hat{D}$ = 297 ± 53 crabs/ha	$\hat{D}$ = 271 ± 49 crabs/ha

Station 67: 52 Crabs

Ungrouped Estimates	Grouped Estimates
$\bar{x}$ = 1.99 m	$\bar{x}$ = 2.07 m
$\hat{\lambda}$ = .510 m^{-1}	$\hat{\lambda}$ = .472 m^{-1}
$\hat{w}$ = 8.12 m^2	$\hat{w}$ = 8.77 m^2
$\hat{D}$ = 285 ± 62 crabs/ha	$\hat{D}$ = 264 ± 58 crabs/ha

TABLE 3: SIGHTING DISTANCES BY STATION

Station 16: 51 Crabs

56 10 38 36 18 51

86 163 122 142 142 142 102 117 173 127 61 173 147 168 112
173 61

254 183 295 224 183 239 254 244 213 254 193 213 244

305 396 335 366 335 396 325 361 335 325

498 503 488 498 498

Station 21: 93 Crabs

10 5 25 20 5 20 51 25 31

81 142 132 147 112 61 76 91 152 165 163 142 71 168 61 152
132 76 61 142 112 142 152 132 132 158 122 158 158

274 305 244 183 239 183 239 264 198 244 290 224 302 234 254
254 254 198 193 264 290 284 183 295 254 183 239 254 193

335 386 386 356 386 335 396 340 356 305 305 345 356 315 330
371 305 356

427 457 427 518 457 457 427 488

Station 67: 52 Crabs

46 41 41 30 41

112 163 107 132 81 122 61 61 168 81 81 112 152 173 163
81

203 185 244 279 274 193 213 193 224 213 193 183 239 284

305 335 305 351 396 366 366 422 386 416 366

427