

Abstract.—The spatial distributions of marine biota are frequently patchy. Samples taken from these populations are characterized by values which are mostly small, relative to the population mean, and a few that are very large. It is therefore difficult to estimate stock size using conventional methods. We performed Monte Carlo simulations based on trawl data for Dungeness crab *Cancer magister* and compared the behavior of three estimators of central tendency: sample mean, geometric mean, and a lognormal estimate. Although the sample mean is unbiased, results indicate that single estimates of the population mean (and thus population estimates obtained using area-swept) may be overly sensitive to extreme values; confidence intervals are large and capture the true value at a level well below that prescribed. Estimates of the geometric mean exhibit more stable behavior about its parameter, with mixed results for the lognormal estimate. We propose a conservative approach based on comparison of trends found in each of the three estimators. Moreover, we suggest that abundance of aggregated stocks should be indexed with an estimator that has more desirable statistical properties, such as the geometric mean. This may reduce error associated with conventional fisheries stock-assessment practices and thus provide for more effective management of overdispersed stocks.

Trawl survey estimation using a comparative approach based on lognormal theory*

Robert A. McConnaughey

School of Fisheries, WH-10
University of Washington, Seattle, Washington 98195
Present address: Alaska Fisheries Science Center, National Marine Fisheries Service,
NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115-0070

Loveday L. Conquest

Center for Quantitative Science, University of Washington
Seattle, Washington 98195

Effective scientific management of fishery resources is dependent upon reliable measures of stock abundance. To this end, research trawl surveys are routinely used in concert with fishery catch statistics to provide estimates of population parameters. The analytical procedures used often rely on the assumption that statistical methods based on normal probability theory are appropriate and, as such, that the individuals comprising the population are not aggregated in space (Elliott 1977). However, marine biota are commonly overdispersed, and frequently it is the *logarithms* of abundance (or biomass) which conform to the normal or Gaussian distribution (reviewed by McConnaughey 1991). Rather than an artifact of sampling, in many cases this spatial attribute is the product of behavioral responses and/or physical processes affecting dispersal (e.g., Epifanio 1987, Dew 1990). Samples taken from these populations are characterized by mostly small values relative to the population mean, and a few very large ones. Under these circumstances, single estimates of the population mean from the arithmetic mean (sample average) may be too low because very large values are of-

ten underrepresented at the levels of sampling effort common to research trawl surveys. When large catches are present in a sample, variance estimates may become excessively high (e.g., Otto 1986). This may introduce a high degree of uncertainty into the resource management process which, if ignored, can have potentially serious repercussions (Ludwig & Walters 1981).

We investigated two alternative measures of central tendency and compared their statistical behavior with that of the arithmetic mean. These alternatives are the geometric mean and a model-based estimate of the arithmetic mean based on lognormal theory. An evaluation of trends based on a comparison of the three estimators is proposed. This approach may identify error associated with the conventional index of abundance, thereby reducing the likelihood of false conclusions concerning trends in stock abundance.

Data and methods

Monthly trawl surveys of Dungeness crab *Cancer magister* abundance along the southern Washington coast provided representative values of density (n/ha) for analysis with Monte Carlo techniques. Density data

such as these are commonly expanded according to area-swept procedures to produce estimates of population size. Both nearshore and estuarine populations were sampled, and we examined data collected during four consecutive years. Survey design and methodology are discussed in Armstrong & Gunderson (1985) and Gunderson et al. (1990).

The spatial dispersion of a population determines the relationship between its mean abundance and variance, and this information may be used to select an advantageous data transformation (Elliott 1977). Strong linear relationships between the means and standard deviations of our density data ($r^2=0.98$ and 0.70 , with $P<0.001$ and $P<0.001$ for the coastal and estuarine areas, respectively) and graphical analysis of log-transformed data suggested a logarithmic transformation would be appropriate. In order to test this assumption, the Kolmogorov test for normality was applied to the density data for each cruise, both before and after transformation (Table 1). These preliminary analyses suggested that the density data were lognormally distributed and, as such, that individuals in the crab population were aggregated in space. However, lognormal theory cannot be applied directly to any

sample that contains a zero value, since the logarithm of zero is undefined. Since our data exhibit only the occasional zero catch, we used the common $\ln(X+1)$ transformation to normalize the data. An alternative approach, when a significant fraction of the data consists of zero catches, would be to use the Δ -distribution (Pennington 1983 and 1986, Smith 1988), which is essentially a lognormal distribution with a proportion (Δ) of zeros.

The lognormal distribution and parameter estimation

The lognormal distribution may be represented as a Gaussian distribution of logarithmic data or, equivalently, as a right-skewed distribution of untransformed data (Aitchison & Brown 1969). A brief review of the density function and relevant parameters for the lognormal distribution appear in the Appendix. There, and throughout the text, we use the following notation to distinguish between untransformed and transformed scales and between population parameters and their estimates: X represents untransformed density values, while \bar{X} (the ordinary sample mean) and s^2_X (the

Table 1

Goodness-of-fit probabilities from Kolmogorov tests for normality with means (\bar{Y}) and standard deviations (s_Y) for log-transformed *Cancer magister* abundance data. Cruise refers to sequential trawl surveys (n =number of samples) in coastal and estuarine areas along the southern Washington coast over a consecutive 4-year period.

Cruise	Coast					Estuary				
	n	Raw ¹	Log ²	\bar{Y}	s_Y	n	Raw	Log	\bar{Y}	s_Y
1	35	.006	.964	4.20	2.35	20	.211	.890	6.96	1.19
2	41	.000	.740	4.64	2.50	20	.129	.816	6.82	1.10
3	42	.000	.687	4.95	2.86	20	.909	.651	6.88	.66
4	38	.000	.738	6.65	2.13	20	.133	.207	6.59	.79
5	42	.000	.932	5.30	2.26	20	.167	.434	6.34	.56
6	44	.014	.287	3.34	1.97	16	.373	.411	4.90	1.72
7	40	.109	.230	3.08	1.72	20	.004	.418	6.04	1.21
8	41	.000	.800	3.59	1.60	20	.088	1.000	5.65	1.21
9	44	.000	.164	2.91	2.08	20	.013	.573	5.68	1.36
10	43	.016	.528	3.90	1.78	20	.140	.878	5.45	.96
11	44	.007	.102	3.16	2.10	20	.172	1.000	4.64	1.16
12	44	.002	.829	3.62	1.93	20	.129	.705	5.81	1.01
13	44	.000	.759	4.55	2.15	20	.435	.979	6.61	1.04
14	44	.000	.507	3.90	2.49	20	.311	.960	5.93	1.06
15	44	.000	.306	5.05	2.20	20	.336	.981	6.42	1.10
16	44	.000	.484	3.62	3.84	20	.027	.931	6.43	1.91
17	44	.000	.471	3.44	4.94	20	.434	.922	6.93	.82
18	44	.000	.370	3.69	4.73	19	.336	.868	6.82	1.18
19	44	.000	.949	4.75	6.89	20	.388	.980	6.43	.92
20	43	.000	.528	3.90	6.48	20	.010	.964	6.23	1.29

¹P-value for untransformed data.

²P-value for log-transformed data.

sample variance) are estimators of μ and σ^2 , the population mean and variance of the untransformed data. Letting $Y = \ln(X)$, \bar{Y} and s^2_Y are estimators of μ_{LN} and σ^2_{LN} , the population mean and variance of the log-transformed data. Note that in Eq. (A3) and (A4) both μ and σ^2 for the lognormal distribution are functions of two parameters, μ_{LN} and σ^2_{LN} , making the former parameters difficult to estimate. In particular, any estimate of μ involves both location and dispersion parameters. Therefore, variation in estimating μ will come from two sources: variation in estimating μ_{LN} and variation in estimating σ^2_{LN} .

The arithmetic mean (AM) The ordinary sample mean is an unbiased estimator of μ regardless of the underlying frequency distribution. When the underlying distribution is normal, the sample mean is also the minimum variance unbiased estimator (MVU, the one with the smallest variance of all unbiased estimators) of μ . However, the sample mean does not have this MVU property when the underlying distribution is lognormal (Gilbert 1987). Moreover, the AM is sensitive to the presence of one or more large data values, particularly for small sample sizes. For lognormal data, these extreme values are not outliers; they simply reflect the right-skewed nature of the distribution. Finney (1941) demonstrated the inefficiency of the sample mean when the variance of the natural logarithms is greater than 0.69, and Koch & Link (1970) suggested using the sample mean only when the coefficient of variation is believed to be less than 120%. For highly-skewed distributions such as the lognormal, sample sizes in excess of 200 may be necessary to invoke the Central Limit Theorem, which justifies use of the sample mean for inferences about means of populations that are not normally distributed (Sissenwine 1978, Jahn 1987).

The Finney-Sichel estimator (FM) Among alternative estimators that have been investigated is an MVU estimator of μ (Finney 1941, Sichel 1952), which also has been described as equivalent to a maximum-likelihood estimator for lognormal data (Aitchison & Brown 1969). The Finney-Sichel method adjusts the geometric mean upwards and is commonly used in gold and trace-mineral assay work, where ore concentrations are typically lognormally distributed (Sichel 1952). If \bar{Y} and s^2_Y represent the ordinary sample mean and variance of the log-transformed values, the Finney-Sichel estimate for μ is

$$FM = \exp(\bar{Y}) \psi_n(t) \tag{1}$$

where n is the sample size and $\psi_n(t)$ is the infinite series

$$1 + \frac{(n-1)t}{n} + \frac{(n-1)^2 t^2}{2!n^2(n+1)} + \frac{(n-1)^3 t^3}{3!n^3(n+1)(n+3)} + \frac{(n-1)^4 t^4}{4!n^4(n+1)(n+3)(n+5)} + \dots \tag{2}$$

with $t = \frac{s^2_Y}{2}$. The function $\psi_n(t)$ is defined such that

$$E[\psi_n(s^2)] = \exp\left(\frac{n-1}{n} \sigma^2\right) \text{ and } \lim_{n \rightarrow \infty} [\psi_n(s^2)] = \exp(\sigma^2);$$

it is used extensively with the lognormal distribution (Smith 1988). In their book, Aitchison & Brown (1969) included tables of ψ_n for computing the Finney-Sichel estimate. More extensive tables are provided in Link et al. (1971), who claim that linear interpolation between tabled values gives close approximation for estimates of μ . They also include a FORTRAN program for calculating the ψ_n function, which we used in computing FM, the Finney-Sichel estimate of the population mean. (A version of this program may be obtained from the authors.)

Confidence limits for the lognormal mean are not symmetric because of the skewed nature of the underlying distribution. Hence, it becomes necessary to compute separate upper and lower confidence limits. Land (1971, 1975) obtained upper one-sided $100(1-\alpha)\%$ and lower one-sided $100\alpha\%$ confidence limits for the lognormal mean, where α is the frequency of type I error:

$$UL_{1-\alpha} = \exp\left(\bar{Y} + \frac{s^2_Y}{2} + \frac{s_Y H_{1-\alpha}}{\sqrt{n-1}}\right) \tag{3}$$

$$LL_{\alpha} = \exp\left(\bar{Y} + \frac{s^2_Y}{2} + s_Y \frac{H_{\alpha}}{\sqrt{n-1}}\right). \tag{4}$$

The quantities $H_{1-\alpha}$ and H_{α} [functions of α , $(n-1)$ and s_Y] are obtained from tables in Land (1975) for sample sizes of $n \geq 3$.

The geometric mean (GM) The geometric mean, $e^{\bar{Y}}$, will be a biased estimate of μ (Appendix) but may be more precise with respect to its population parameter than will be the case for estimators of the population mean. (Actually, $E(e^{\bar{Y}}) = \exp(\mu_{LN} + \frac{1}{2n\sigma^2_{LN}})$ so the GM is

biased even for $e^{\mu_{LN}}$, but this bias decreases rapidly as n increases.) When exponentiated, the population mean of the transformed data, μ_{LN} , is the geometric mean catch and, equivalently, the median catch for lognormal data. It remains unaffected by skewness, a function of $[\exp(\sigma^2_{LN}-1)]$. It is less affected by large values of X , owing to the nature of a logarithmic transformation; hence, its sampling distribution is less skewed

than that for the AM. Aitchison & Brown (1969) note that "since the arithmetic mean involves both the location and dispersion parameters, it is not a pure measure of the [response variable] under the lognormal hypothesis: for this the geometric mean or median is to be preferred."

Monte Carlo simulations based on crab trawl data

Monte Carlo simulations consist of calculations made on data sets whose elements are randomly selected from specified probability distributions. This approach permits an evaluation of various point-estimation procedures on the basis of expected outcomes. It also allows a closer examination of individual cases than is possible with a purely analytical approach and permits an evaluation of the effects of sample size. For this investigation, single values of mean density and standard deviation were calculated for each cruise in the two trawl locations along the Washington coast. The means of these statistics were used to define two representative lognormal distributions, which are identified as lognormal (4,2) for the coastal area and lognormal (6,1) for the estuarine area. These distributions have means of 4.0 and 6.0, and standard deviations of 2.0 and 1.0, respectively, for the log-transformed variable; they will be referred to as LOGN (4,2) and LOGN (6,1) (Fig. 1). From these two probability distributions, we created 1000 sets of simu-

lated density data for each of 13 sample sizes (2,4,6,8,10,15,20,25,30,35,40,45,50) using a pseudorandom number generator (Minitab, Inc., University Park PA). Sample sizes were selected to encompass the range of values associated with ongoing trawl surveys. Table 2 presents descriptive statistics for each of these data sets. (These data sets are archived on magnetic tape, and access can be arranged through the authors.)

We investigated three methods of estimating central tendency. The AM method consisted of computing arithmetic means and traditional confidence intervals (e.g., at 90% confidence) based on the Student's *t*-distribution. The FM method used the Finney-Sichel estimator for the mean of a lognormal distribution as presented in Eq. (1) and (2). For confidence limits, the method by Land (1971, 1975) as presented in Eq. (3) and (4) was used. The GM method used $e^{\bar{y}}$ as an estimate of $e^{\mu_{LN}}$, the geometric mean (or median) in the untransformed scale. A 90% confidence interval was derived as follows:

$$\exp \left(\bar{Y} - t_{n-1} \frac{s_Y}{\sqrt{n}} \right), \exp \left(\bar{Y} + t_{n-1} \frac{s_Y}{\sqrt{n}} \right). \quad (5)$$

This method estimates a different parameter (the median rather than the population mean) than the first two methods. However, because the median is asymptotically a function of only a single parameter, μ_{LN} , the GM method tends to give more stable estimates of its parameter, and it is worthwhile to compare its performance as another index of central tendency to the first two methods.

Comparison measures to evaluate performance of the estimators

We used the following measures of comparison to evaluate the performance of the estimators: root mean squared error (RMSE), deviation of the estimate from the true parameter (BIAS), average length of the 90% confidence interval (AVL), standard deviation of the 90% confidence interval length (SDCI), and percent containment of the parameter by the confidence interval estimate (PERCON). These were estimated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{1000} (\text{estimated parameter from } i^{\text{th}} \text{ data set} - \text{true value})^2}{1000}}$$

Root mean squared error is a measure of the average variation in the estimated mean relative to the true

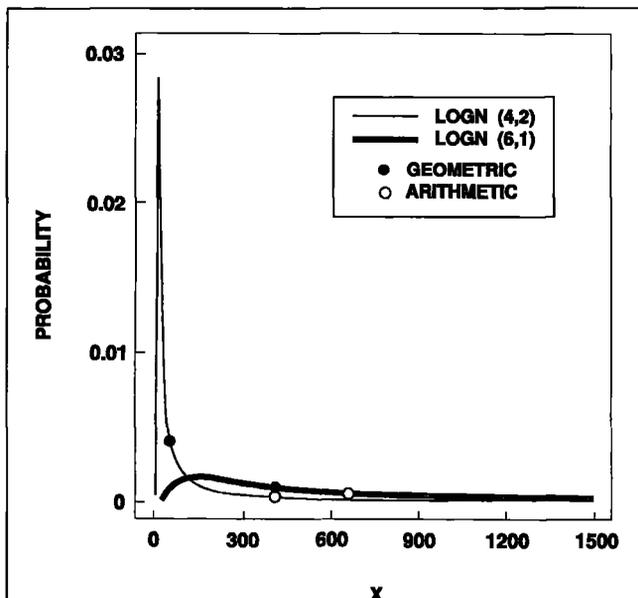


Figure 1

Probability density functions for the representative lognormal distributions. Included are reference marks to indicate values of the respective arithmetic and geometric means.

Table 2a

Descriptive statistics for 1000 simulated trawl data sets representative of **coastal** populations of *Cancer magister* in Washington [Lognormal (4,2)]. Trimmed means calculated using the central 90% of the individual data sets. Min/Max refer to the minimum and maximum values in the data set.

n	Arithmetic mean		Trimmed mean		Standard deviation		Min	Max
	Raw ^a	Log ^b	Raw	Log	Raw	Log		
2	431	4.104	192	4.106	1,949	2.026	<1	48,860
4	382	4.003	163	4.004	1,714	1.999	<1	64,248
6	398	3.987	162	3.988	1.946	2.007	<1	62,131
8	370	3.985	157	3.984	1,818	1.982	<1	71,789
10	386	4.001	161	4.002	1,964	1.996	<1	110,761
15	378	4.002	163	4.007	1,723	2.002	<1	82,492
20	391	4.004	162	4.003	1,920	1.997	<1	122,967
25	394	3.999	159	4.002	2,313	1.991	<1	160,546
30	407	4.002	161	4.004	2,603	2.000	<1	236,341
35	422	4.016	162	4.015	3,301	1.992	<1	380,743
40	419	3.999	161	3.999	7,609	1.997	<1	1,479,353
45	397	3.989	162	3.989	2,167	2.010	<1	226,970
50	411	4.003	163	4.003	2,525	2.008	<1	323,734

True value of μ is $403.43 = \exp\left(4 + \frac{2^2}{2}\right)$; (see Appendix).

^a untransformed density (n/ha).

^b log-transformed density (n/ha).

Table 2b

Descriptive statistics for 1000 simulated trawl data sets representative of **estuarine** populations of Dungeness crab *Cancer magister* in Washington [Lognormal (6,1)]. Trimmed means calculated using the central 90% of the individual data sets. Min/Max refer to minimum and maximum values in the data set.

n	Arithmetic mean		Trimmed mean		Standard deviation		Min	Max
	Raw ^a	Log ^b	Raw	Log	Raw	Log		
2	633	5.968	528	5.971	739	.994	10	5,703
4	676	6.007	552	6.007	860	1.012	10	10,525
6	682	6.026	557	6.024	924	.993	11	24,259
8	690	6.010	553	6.006	955	1.016	12	20,499
10	659	5.993	541	5.995	835	1.004	6	13,333
15	650	5.986	537	5.986	816	.993	12	18,877
20	655	5.992	538	5.993	825	.998	8	13,756
25	667	6.000	544	5.999	868	1.003	4	19,785
30	664	5.995	543	5.995	857	1.004	5	21,060
35	664	5.999	541	5.998	901	.995	7	30,309
40	657	5.996	540	5.995	844	.993	9	20,555
45	666	6.004	547	6.004	869	.997	6	30,655
50	664	6.001	544	6.002	856	.998	7	18,725

True value of μ is $665.14 = \exp\left(6 + \frac{1^2}{2}\right)$; (see Appendix).

^a untransformed density (n/ha).

^b log-transformed density (n/ha).

mean density and, as such, is a measure of accuracy. For any unbiased estimator (e.g., the AM or FM, where the expected value of the estimator is the parameter itself), the RMSE is the same as the variance of the estimator, in terms of expected value. For a biased estimate (recall that GM has positive bias in estimating $e^{\mu_{LN}}$, which decreases as n increases), RMSE incorporates both bias and variance.

$$BIAS = \frac{\sum_{i=1}^{1000} (\text{estimated parameter from } i^{\text{th}} \text{ data set} - \text{true value})}{1000}$$

= (average value of estimated parameter—true parameter).

The bias is the average amount by which the estimate tends to “miss” its respective parameter.

$$AVL = \frac{\sum_{i=1}^{1000} (UL - LL)_i}{1000} = \frac{\sum_{i=1}^{1000} \text{length}_i}{1000}$$

where $(UL-LL)_i$ = length of a single 90% confidence interval for the i^{th} data set. The average length is a measure of precision.

$$SDCI = \sqrt{\frac{\sum_{i=1}^{1000} (\text{length}_i - AVL)^2}{1000 - 1}}$$

The standard deviation is a measure of the spread of the confidence-interval lengths around the average length. An estimator with the most reproducible estimate of the precision of the estimated mean would have confidence intervals of relatively low variability in length.

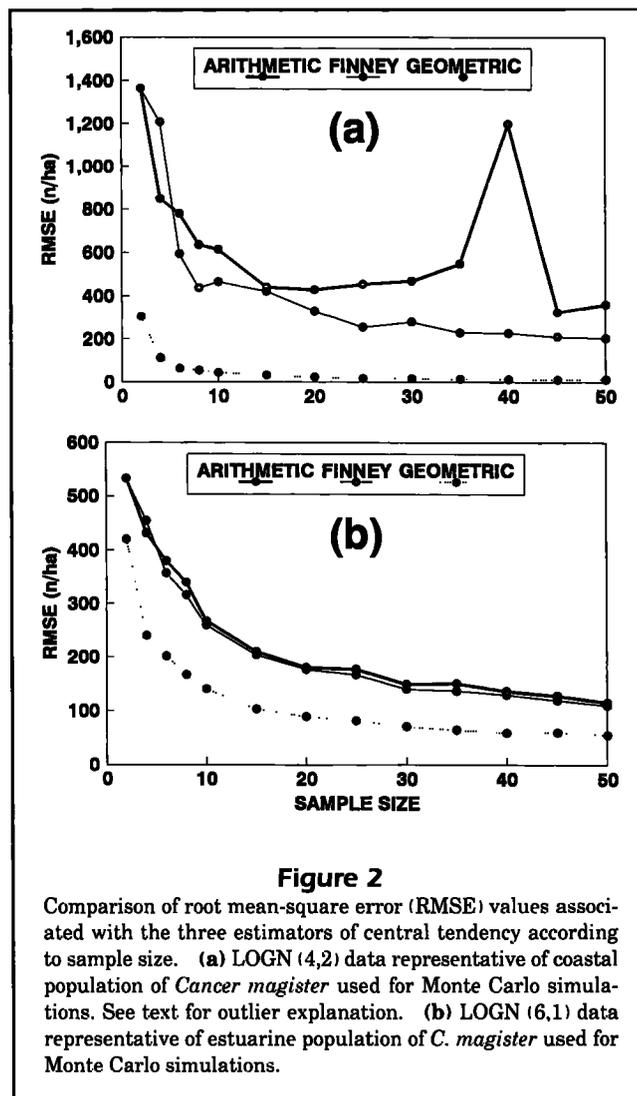
The frequency with which a confidence interval includes the true value of the parameter defines the containment rate, PERCON. If the assumptions of sampling and the appropriateness of statistical model are met, 90% confidence intervals should contain the density parameter being estimated approximately 90% of the time.

The three estimators of central tendency (AM, FM, GM) and their confidence intervals were also calculated for actual density data obtained during the monthly trawl surveys. Two large systems, termed the Coast and the Estuary, were considered.

Results

Monte Carlo simulations

Root mean squared error The RMSE was consistently lower for the GM than for the other measures of central tendency (Fig. 2). The FM provided point estimates of μ that were consistently more accurate (except at very small sample sizes) than the AM method, particularly as skewness of the density data (Fig. 1) increased. The RMSE of GM estimates and of μ obtained with the FM declined steadily as sample size increased, whereas that for the AM, although generally declining, was somewhat less regular and much more erratic (see Fig. 2a, $n=40$). Closer inspection of the LOGN (4,2) data set revealed a single extreme

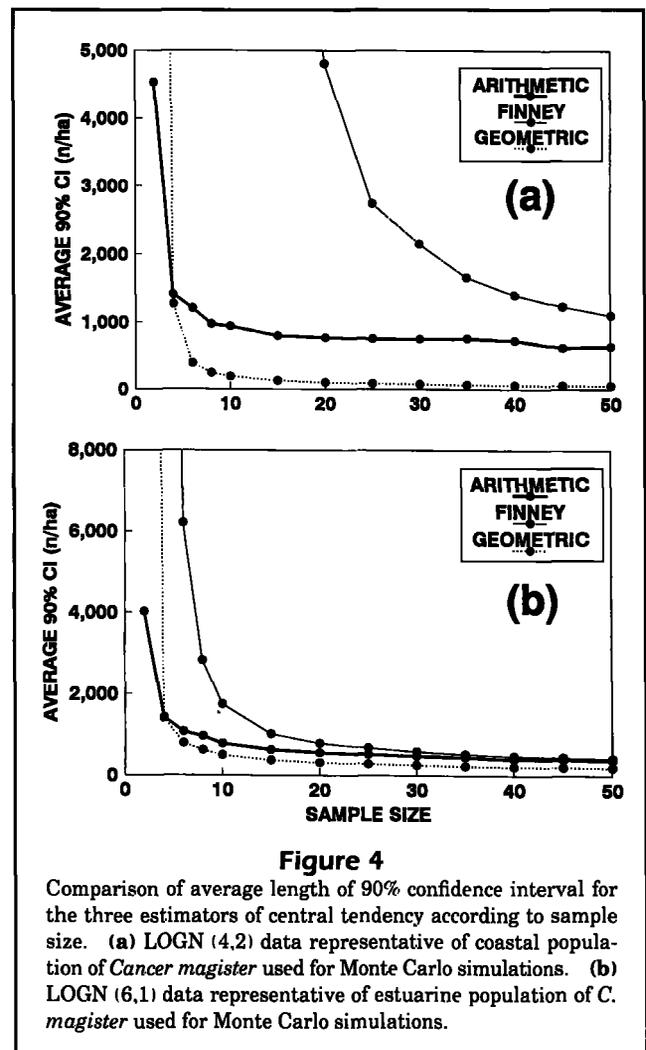
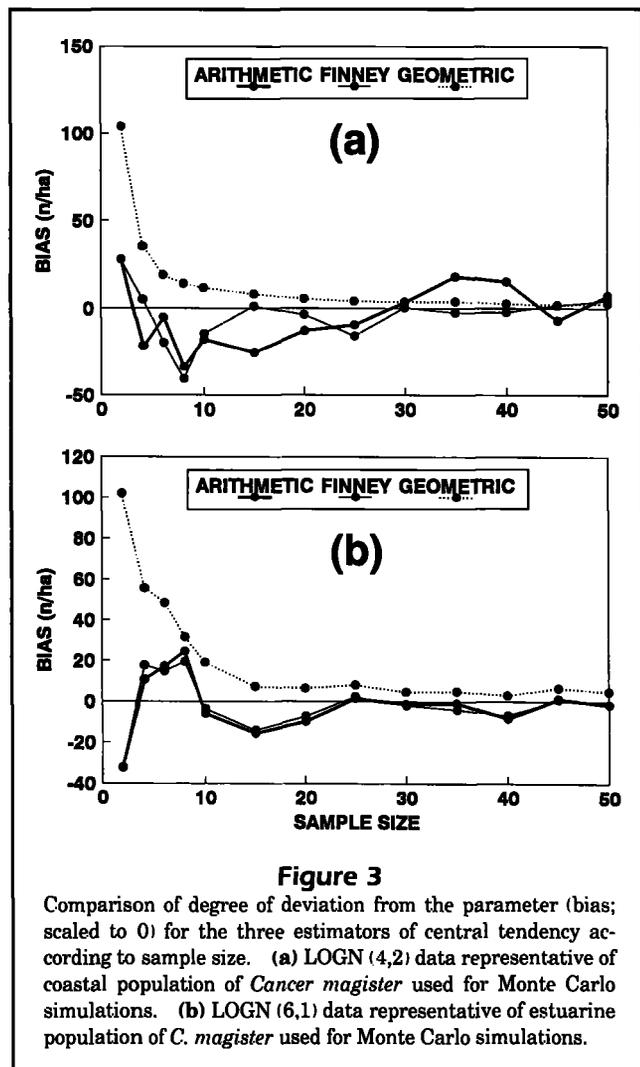


value out of 40,000 data points that caused a considerable increase in the RMSE associated with the AM estimate. It is noteworthy that the magnitude of this simulated density value is in keeping with extreme values observed in the field. Accuracy of GM estimation improved dramatically as skewness increased, in contrast to the FM and AM responses wherein accuracy decreased as skewness increased.

Deviation of the estimator from the parameter (bias) Overall, the most extreme deviations were associated with the smallest sample sizes; this disparity decreased as sample sizes increased (Fig. 3). GM estimates deviated less, stabilized at smaller sample sizes, and, despite the positive bias, converged much more predictably to $e^{\mu LN}$ than did AM and FM in estimating μ . In general, estimates of μ oscillated about the parametric value and converged as sample size increased. The absolute values of the deviations from μ are smaller

for the FM than for the AM method in 17 of the 26 cases examined, without an obvious trend related to the skewness of the data. It is noteworthy that estimates of μ obtained with the AM and FM methods are equivalent when $n=2$.

Average length of the interval estimate The average length of the 1000 90% confidence intervals (CIs) calculated for each sample size was consistently shorter for the GM (which only estimates one parameter) than for the intervals of the AM or FM (Fig. 4). Intervals calculated using the FM method were consistently larger than those for the AM method. Overall, the degree of difference between the three estimators decreased as sample size increased and as skewness decreased. Average lengths were inordinately large at the smallest sample sizes and decreased rapidly thereafter. The average CI length for the GM decreased as skewness increased, in contrast with the behavior of CI lengths for μ .



Standard deviation of the confidence interval length The standard deviation of the 1000 90% CIs calculated for each sample size was consistently lower for the GM than for the AM or FM, both in an absolute sense and relative to the average CI length (Fig. 5). At smaller sample sizes, the standard deviations for the AM were less than those for the FM. However, this pattern was reversed at larger sample sizes such that the FM had the more precise interval estimates (note the crossovers at $n=35$ and $n=25$ in Figs. 5a and 5b, respectively). Overall, the precision of the interval estimates declined as sample size decreased and as skewness increased; the effect was most pronounced for the FM method. The GM response was unique in that precision increased as skewness increased. Of particular note is the dramatic loss of precision of the AM interval estimate apparent in Fig. 5a ($n=40$) which, upon investigation, was attributed to a single extreme value.

Parameter containment within the interval estimate The GM parameter $e^{\mu LN}$ occurred within its interval estimates, as did μ within the intervals obtained by using FM, at the prescribed 90% confidence level (Fig. 6). The rate of GM containment oscillated within 1–2% of this level under all circumstances. Confidence intervals for the FM contained μ at the rates of 89.1–92.3% (Fig. 6a) and 89.1–93.6% (Fig. 6b); the highest percentages were associated with the smallest sample size, perhaps due to their relatively greater lengths (Fig. 5). In contrast, AM interval estimates contained μ at rates of 47.7–65.8% (Fig. 6a) and 76.9–85.3% (Fig. 6b), well below the prescribed level of confidence.

Dungeness crab trawl survey data

We also computed the three estimators for actual *C. magister* density data to assess the gain in informa-

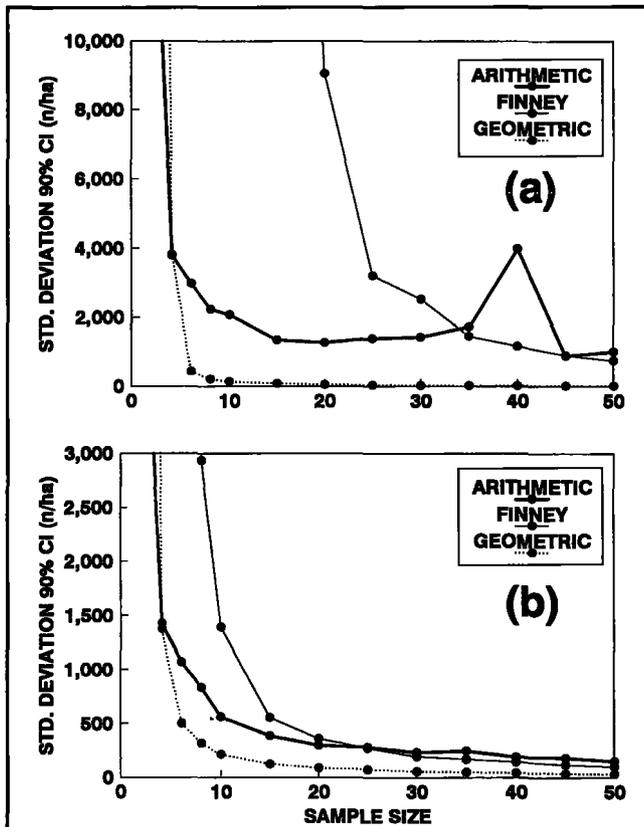


Figure 5

Comparison of standard deviation of the length of 90% confidence interval for the three estimators of central tendency according to sample size. (a) LOGN (4,2) data representative of coastal population of *Cancer magister* used for Monte Carlo simulations. (b) LOGN (6,1) data representative of estuarine population of *C. magister* used for Monte Carlo simulations.

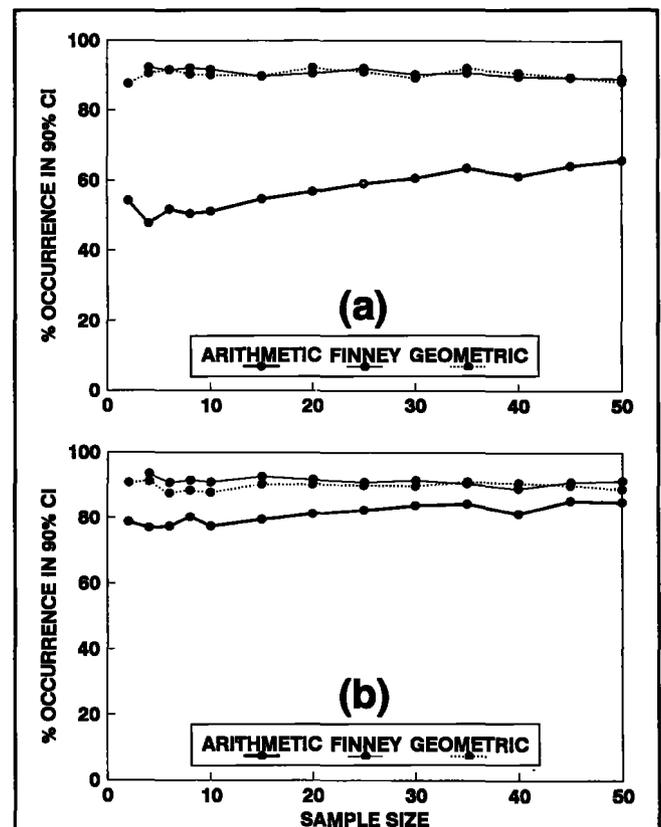


Figure 6

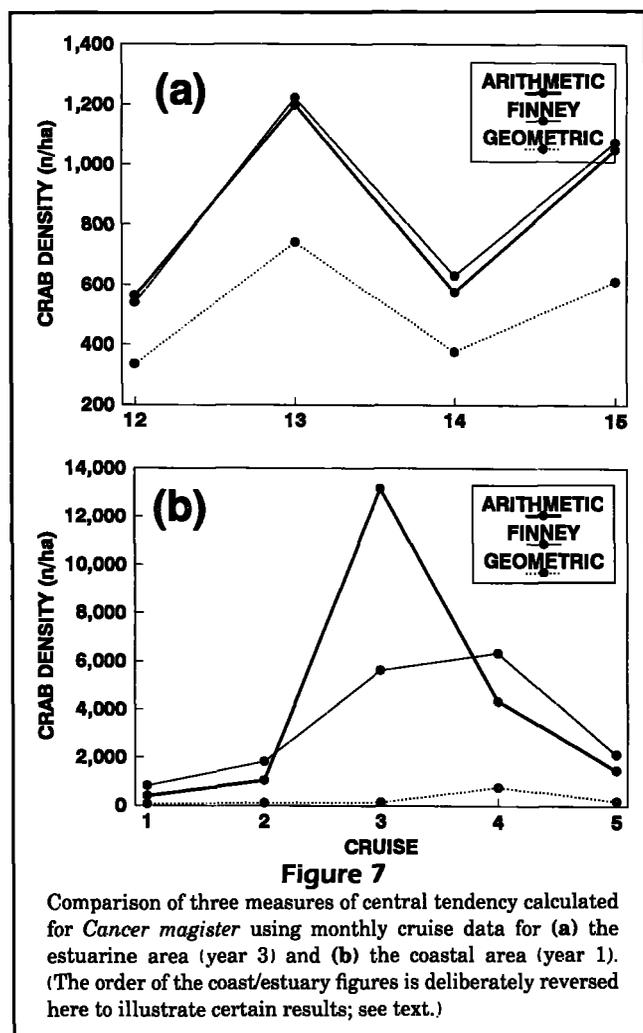
Comparison of percent occurrence of the parameter in the 90% confidence interval for the three estimators of central tendency according to sample size of *Cancer magister*. (a) LOGN (4,2) data representative of coastal population of *C. magister* used for Monte Carlo simulations. (b) LOGN (6,1) data representative of estuarine population of *C. magister* used for Monte Carlo simulations.

tion over using any single estimator alone. For the Estuary data, trends in abundance routinely paralleled one another, differing only by their relative magnitude (Fig. 7a). Characteristically, estimates of μ from the survey data obtained with the FM method exceeded those of the AM, which, in turn, exceeded estimates of the GM parameter ($e^{\mu_{LN}}$). In some cases, trends in the Coast estimates were diametrically opposed (Fig. 7b). As expected, the GM estimate was consistently lower than both AM and FM, reflecting the difference in the population parameter being estimated. Noteworthy was the reversal in the relative magnitudes of the AM and FM estimates during the interval between Cruise 2 and Cruise 4.

Discussion

Conventional analysis of catch data and alternatives

Population estimates are routinely generated using untransformed catch data and arithmetic mean calcu-



lations (e.g., BIOMASS procedure of the U.S. NMFS, Gunderson et al. 1978; STRAP procedure of Can. Dep. Fish. & Oceans, Smith & Somerton 1981). Several methods for reducing the variance associated with these estimates of abundance have been used, often despite recognizable limitations. These fall broadly into two categories: (1) model-based approaches, which model the underlying distribution of the data, and (2) design-based approaches, which rely upon probability sampling and large sample results. Smith (1990) compared the two approaches for estimating resource abundance with trawl surveys and concluded with an example of a model-based predictive estimate using additional information (salinity, temperature, depth). Other examples of model-based estimation in fisheries applications include use of a weighted negative binomial distribution (Zweifel & Smith 1981), the delta distribution (Pennington 1983), and the geostatistical technique of kriging (Conan 1985). Stratification of the sampling frame is a common example of a design-based approach. Although this is theoretically appealing, Gavaris & Smith (1987) have demonstrated that stratified random sampling may be inferior to a simple random design, because of suboptimal allocation of stations to strata. They suggest that a decrease in the number of strata used in the eastern Scotian Shelf groundfish survey would provide for more flexible allocation of total sampling effort in the future. Unfortunately, many of the problems attendant with specifying stratum boundaries will persist; these include interannual variability in distribution and abundance of stocks related to environmental factors and the typical multi-species scope of most research trawl surveys. Because of these difficulties, catch data are commonly stratified *after* sampling is completed (Picquelle & Stauffer 1985, Otto 1986). However, post-stratification (i.e., *a priori* examination of catch data for the purpose of assigning strata) is not a valid approach and is not recommended (Cochran 1977).

Other methods for estimation of abundance are expedient, yet may be based on the specious assumption that extreme values are "outliers" and are therefore not integral to the data set. Included is the practice of eliminating extreme values or the use of trimmed (or Winsorized) means (Halliday & Koeller 1981, Bates 1987, Harding et al. 1987, Smith 1981). Ignoring instances where human error is involved, these *ad hoc* procedures may introduce substantial negative bias to estimates of the true population mean (compare μ and the trimmed means in Table 2), thereby contributing to misleading conclusions about trends in the data.

Design-based and model-based approaches

A strict probability sampling approach (i.e., design-based and without any underlying models) requires

that resulting estimates be normally distributed according to the Central Limit Theorem. However, Hansen et al. (1983) state, "When surveys use relatively small samples, the samples may be too small for the application of the theory [for large samples] to be essentially assumption-free." Our approach is model-based. As long as one is restricted to samples that may not be considered "acceptably large" (and further hampered by considerable skewness caused by extreme values), use of a model-based approach is not unwarranted (Little 1983).

With regard to robustness, Myers & Pepin (1990) argue that exclusive use of a lognormally-based estimate can be sensitive to model assumptions, leading to possible bias and reduction in efficiency. Because contamination of a lognormal distribution with data from similarly-shaped distributions (e.g., Weibull or gamma) is difficult to detect for sample sizes less than 40, they suggest using lognormally-based estimators of abundance only when there is evidence that the underlying population is lognormal. Obviously, use of transformations and model-based estimators is a "double-edged sword," and these procedures should not be applied indiscriminately. When appropriate (e.g., Table 1), however, significant improvement in the relative efficiency of the sample average and, in particular, the estimated variance, can be realized (e.g., Finney 1941, Koch & Link 1970, Myers & Pepin 1990).

A comparative approach

If nothing is known about the spatial distribution of an organism, the sampling plan must be designed to determine distribution patterns as well as population size. Knowledge of the distribution pattern aids in selection of the proper estimation procedure. Based on the arguments presented above, combined with the rather ubiquitous nature of overdispersion in the marine environment, we prefer an approach based on three estimators, namely the arithmetic mean, the geometric mean, and the Finney-Sichel estimator of μ . By taking a comparative approach, one may be reasonably certain of apparent trends in the data if the trend is consistent for the three estimators. For the estuarine crab population illustrated in Figure 7a, the parallel behavior of the estimates corresponded to changing values of \bar{Y} coupled with nominal changes in s^2_Y (Table 1). In this case, there is no evidence to suggest that conventional analysis of catch data (i.e., using the AM method) was less than adequate. However, trends in the estimates may, on occasion, be opposed to one another, as was demonstrated for the coastal crab population (Fig. 7b). The AMs suggest a precipitous drop in abundance occurred during the interval between

Cruises 3 (with two extreme values) and 4, whereas both the FM and GM procedures indicated a moderate increasing trend during the same period. From the behavior of the three estimators, we conclude that between Cruises 2 and 4, μ_{LN} (and μ) may have increased slightly, but σ^2_{LN} (and thus the skewness of the distribution) probably increased and then decreased, affecting the FM and AM estimates (the latter more strongly) but not the GM estimate. This is verified by checking the \bar{Y} and s^2_Y values in Table 1. Changes in skewness relate directly to the size of the larger catches and, hence, the degree of spatial aggregation in the population. Plotting the three estimators and relating the trends back to changes in \bar{Y} and changing s^2_Y has yielded some insight into the behavior of the estimators. It has also allowed us to extract more information about the crab population than if we had used only one estimator, the AM. In cases such as this, where there is significant disagreement among the estimators, the data set should be carefully evaluated as to its underlying probability distribution and the most appropriate index selected. If the lognormal distribution is reasonable, the GM may well be the preferred estimator (Aitchison & Brown 1969); use of the GM may be advantageous because it is relatively insensitive to extreme values (particularly so for highly-skewed data) in terms of accuracy and precision. Since catch coefficients are not routinely considered with trawl-survey data of this type, the resulting stock-size estimates are, strictly speaking, *indices* of abundance (Caddy 1986). Under these circumstances, it may be advantageous to use an alternative estimator of central tendency, such as the GM, to generate the index.

Acknowledgments

We thank Drs. David Armstrong and Donald Gunderson of the University of Washington School of Fisheries for access to unpublished data and for their comments and suggestions. Their research program was supported by an institutional grant from Washington Sea Grant (#NA86AA-D-SG044 Project R/F-68) and the U.S. Army Corps of Engineers (#DACW67-85-C-0033). In addition, we thank the following for helpful comments and suggestions on previous drafts of this paper: Dr. Robert Crittenden (University of Washington, School of Fisheries), Charles Douglas Knechtel (FishStat Statistical Helps Service, Seattle), Stephen J. Smith (Marine Fish Division, Bedford Institute of Oceanography), and two anonymous reviewers.

Citations

- Aitchison, J., & J.A.C. Brown**
 1969 The lognormal distribution, with special reference to its uses in economics. Cambridge Univ. Press, Cambridge, 176 p.
- Armstrong, D.A., & D.R. Gunderson**
 1985 The role of estuaries in Dungeness crab early life history: A case study in Grays Harbor, Washington. In Melteff, B.R. (ed), Proceedings of the symposium on Dungeness crab biology and management, p.145-170. Alaska Sea Grant Prog. Rep. 85-3, Univ. Alaska, Anchorage.
- Bates, R.D.**
 1987 Estimation of egg production, spawner biomass and egg mortality for walleye pollock, *Theragra chalcogramma*, in Shelikof Strait from ichthyoplankton surveys during 1981. NWAFC Proc. Rep. 87-20, NMFS Alaska Fish. Sci. Cent., Seattle, 192 p.
- Caddy, J.F.**
 1986 Stock assessment in data limited situations: The experience in tropical fisheries and its possible relevance to evaluation of invertebrate resources. In Jamieson, G.S., & N. Bourne (eds.), North Pacific workshop on stock assessment and management of invertebrates, p. 379-392. Can. Spec. Publ. Fish. Aquat. Sci. 92.
- Cochran, W.G.**
 1977 Sampling techniques. John Wiley, NY, 428 p.
- Conan, G.Y.**
 1985 Assessment of shellfish stocks by geostatistical techniques. ICES (Int. Counc. Explor. Sea) CM/K 30, 24 p.
- Dew, C.B.**
 1990 Behavioral ecology of podding red king crab, *Paralithodes camtschatica*. Can. J. Fish. Aquat. Sci. 47:1944-1958.
- Elliott, J.M.**
 1977 Statistical analysis of samples of benthic invertebrates. Freshwater Biol. Assoc. Sci. Publ. 25, 157 p.
- Epifanio, C.E.**
 1987 The role of tidal fronts in maintaining patches of brachyuran zoeae in estuarine waters. J. Crustacean Biol. 7:513-517.
- Finney, D.J.**
 1941 On the distribution of a variate whose logarithm is normally distributed. J. R. Stat. Soc. Suppl. 7:155-161.
- Gavaris, S., & S.J. Smith**
 1987 Effect of allocation and stratification strategies on precision of survey abundance estimates for Atlantic cod (*Gadus morhua*) on the eastern Scotian Shelf. J. Northwest Atl. Fish. Sci. 7:137-144.
- Gilbert, R.O.**
 1987 Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, NY, 320 p.
- Gunderson, D., D. Worlund, & B. Gibbs**
 1978 Users guide to BMS09: Estimation of biomass and population size by the area-swept method. NWAFC internal doc., NMFS Northwest Fish. Sci. Cent., Seattle 98112, 23 p.
- Gunderson, D.G., D.A. Armstrong, Y. Shi, & R.A. McConnaughey**
 1990 Patterns of estuarine use by juvenile English sole (*Parophrys vetulus*) and Dungeness crab (*Cancer magister*). Estuaries 13:59-71.
- Halliday, R.G., & P.A. Koeller**
 1981 A history of Canadian groundfish trawling surveys and data usage in ICNAF divisions 4TVWX. In Doubleday, W.G., & D. Rivard (eds.), Bottom trawl surveys, p. 27-41. Can. Spec. Publ. Fish. Aquat. Sci. 58.
- Hansen, M.H., W.G. Madow, & B.J. Tepping**
 1983 An evaluation of model-dependent and probability-sampling inferences in sample surveys. J. Am. Stat. Assoc. 78:776-793.
- Harding, G.C., J.D. Pringle, W.P. Vass, S. Pearre Jr., & S.J. Smith**
 1987 Vertical distribution and daily movements of larval lobsters *Homarus americanus* over Browns Bank, Nova Scotia. Mar. Ecol. Prog. Ser. 41:29-41.
- Jahn, A.E.**
 1987 Precision of estimates of abundance of coastal fish larvae. In Hoyt, R.D. (ed.), 10th annual larval fish conference, p. 30-38. Am. Fish. Soc. Symp. 2.
- Koch, G.S. Jr., & R.F. Link**
 1970 Statistical analysis of geological data. John Wiley, NY, 375 p.
- Land, C.E.**
 1971 Confidence intervals for linear functions of the normal mean and variance. Ann. Math. Stat. 42:1187-1205.
 1975 Tables of confidence limits for linear functions of the normal mean and variance. Sel. Tables Math. Stat. 3:385-419.
- Link, R.F., G.S. Koch Jr., & J.H. Schuenemeyer**
 1971 Statistical analysis of gold assay and other trace-element data. U.S. Bur. Mines Rep. Invest. 7495, 127 p.
- Little, R.A.**
 1983 Comment on: An evaluation of model-dependent and probability sampling inferences in sample surveys. J. Am. Stat. Assoc. 78:797-799.
- Ludwig, D., & C.J. Walters**
 1981 Measurement errors and uncertainty in parameter estimates for stock and recruitment. Can. J. Fish. Aquat. Sci. 38:711-720.
- McConnaughey, R.A.**
 1991 Factors affecting Dungeness crab (*Cancer magister*) year class strength along the Washington coast. Ph.D. diss., School Fish., Univ. Wash., Seattle, 162 p.
- Myers, R.A., & P. Pepin**
 1990 The robustness of log-normal estimators of abundance. Biometrics 46:1185-1192.
- Otto, R.S.**
 1986 Management and assessment of eastern Bering Sea king crab stocks. In Jamieson, G.S., & N. Bourne

(eds.), North Pacific workshop on stock assessment and management of invertebrates, p. 83–106. Can. Spec. Publ. Fish. Aquat. Sci. 92.

Pennington, M.

1983 Efficient estimators of abundance for fish and plankton surveys. *Biometrics* 39:281–286.

1986 Some statistical techniques for estimating abundance indices from trawl surveys. *Fish. Bull., U.S.* 84:519–525.

Picquelle, S., & G. Stauffer

1985 Parameter estimation for an egg production method of northern anchovy biomass estimation. In Lasker, R. (ed.), An egg production method for estimating spawning biomass of pelagic fish: Application to the northern anchovy, *Engraulis mordax*, p. 7–15. NOAA Tech. Rep. NMFS 36.

Sichel, H.S.

1952 New methods in the statistical evaluation of mine sampling data. *Trans. Inst. Mining Metal. (Lond.)* 61:261–288.

Sissenwine, M.P.

1978 Using the USA research vessel spring bottom trawl survey as an index of Atlantic mackerel

abundance. *Int. Comm. Northwest Atl. Fish. Sel. Pap.* 3:49–55.

Smith, S.J.

1981 A comparison of estimators of location for skewed populations, with applications to groundfish trawl surveys. In Doubleday, W.G., & D. Rivard (eds.), Bottom trawl surveys, p. 154–163. Can. Spec. Publ. Fish. Aquat. Sci. 58.

1988 Evaluating the efficiency of the Δ -distribution mean estimator. *Biometrics* 44:485–493.

1990 Use of statistical models for the estimation of abundance from groundfish trawl survey data. *Can. J. Fish. Aquat. Sci.* 47:894–903.

Smith, S.J., & G.D. Somerton

1981 STRAP: A user-oriented computer analysis system for groundfish research trawl survey data. Can. Tech. Rep. Fish. Aquat. Sci. 1030, 66 p.

Zweifel, J.R., & P.E. Smith

1981 Estimation of abundance and mortality of larval anchovies (1951–75): Application of a new method. *Rapp. P.-V. Réun. Cons. Int. Explor. Mer* 178:248–259.

Appendix

A random variable X is considered to be lognormally distributed when the natural logarithm of X , $Y=\ln(X)$, has a normal distribution. Specifically, if Y is normally distributed with mean μ_{LN} and standard deviation σ_{LN} , then $X=e^Y$ is lognormally distributed with density function (Aitchison & Brown 1969):

$$f(X) = \frac{1}{\sqrt{2\pi} \sigma_{LN} X} \exp \left(\frac{-(\ln X - \mu_{LN})^2}{2\sigma_{LN}^2} \right). \quad (A1)$$

The k th moment about zero, $E(X^k)$ is expressed as

$$E(X^k) = E(e^{kY}) = \exp \left(k\mu_{LN} + \frac{\sigma_{LN}^2 k^2}{2} \right). \quad (A2)$$

In particular,

$$\text{Population Mean} = \mu = \exp \left(\mu_{LN} + \frac{\sigma_{LN}^2}{2} \right), \quad (A3)$$

$$\text{Variance } (X) = \sigma^2 = [\exp(\sigma_{LN}^2) - 1] \cdot [\exp(2\mu_{LN} + \sigma_{LN}^2)], \quad (A4)$$

$$\text{Geometric Mean} = \text{Median } (X) = \exp(\mu_{LN}). \quad (A5)$$