SYSTEMATIC SAMPLING IN A PLANKTONIC ECOSYSTEM

E. L. VENRICK[1]

ABSTRACT

Two sampling studies, computer simulation and field, investigated the consequences of applying restricted systematic sampling (at predetermined depths) to estimate total chlorophyll in the water column. Comparison was made with stratified random designs with one and two samples per strata. Systematic sampling appeared more accurate than most stratified random designs. However, when repeated over restricted spatial or temporal intervals, systematic designs tended to produce biased estimates. In the central Pacific, an interval of several days, or 100-200 km, appeared necessary for natural population fluctuations to average out the bias inherent in a restricted systematic sampling design.

Underlying sampling theory is the assumption of random collection of samples. This is the only satisfactory method of assuring a representative sample from an unknown population. In pelagic ecology (and undoubtedly in other fields) this assumption is generally neglected and surveys are conducted at fixed geographic positions, at fixed spatial or temporal intervals, and/or at fixed depths, without recourse to randomization. The implicit assumption is that the natural complex variability of pelagic populations provides the necessary element of randomization.

Two types of sampling strategies are frequently called systematic. The present study is concerned with the situation in which the sampling positions are fixed according to some pattern determined by the investigator and are not necessarily at equal intervals; this will be termed restricted systematic sampling (RSS) to distinguish it from the strategy in which only the sampling interval is fixed and the location of the first sample in the first interval is determined at random (randomly located systematic sampling; Yates 1948). Among the alternate sampling strategies which provide the requisite randomization, unrestricted random and stratified random sampling (SR) have received the most attention. In unrestricted random sampling, samples are selected individually from the entire population by some random process, such as by numbering all sampling units and selecting from them by means of a random numbers table. In SR, the population is first divided into subpopulations from each of which one or more samples are selected at random. SR is useful because it ensures that the samples are distributed throughout the entire population.

Three characteristics of sampling designs are of interest (Figure 1): 1) bias, any consistent deviation between the true population parameter and repeated estimates based on the same sampling design; 2) precision, the variability of successive estimates about their mean when a sampling design is repeated on the same population; and 3)

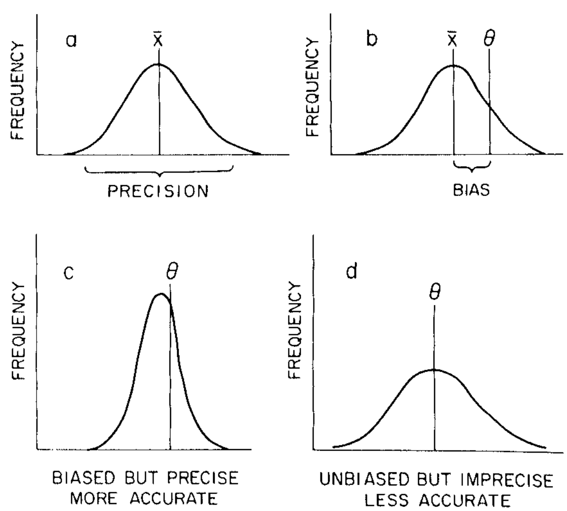

FIGURE 1.—Normal frequency distributions used to illustrate: a) precision, the spread of observations about their mean value ($\bar{x}$); b) bias, the deviation of the mean of repeated observations from the true parameter (θ); c) a distribution which is biased but precise; and d) a distribution which is unbiased but imprecise. Distribution c will be more accurate than distribution d, in spite of the bias, if the average deviation of observations from θ is smaller.

[1]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093.

Manuscript accepted February 1978.
FISHERY BULLETIN: VOL. 76, NO. 3, 1978.

accuracy, a concept including both freedom from bias and high precision and which, in the absence of bias, is equivalent to precision. The practical determination of precision, in its strictest sense, is restricted to quasi-static populations in which the population remains unchanged between collections of replicate samples (forests or soil types or mussel beds, etc.). In the case of RSS, the concept of precision has no meaning in this type of population because successive application of the same sampling design to the same population will give identical results. Such static populations do not exist in a planktonic system because spatial and temporal variability produce continual change. Thus, in the present study, the concept of a population is expanded to incorporate spatial and temporal fluctuations in which case the precision of RSS has a real value.

Theoretical aspects of systematic random sampling strategies have been considered by many (e.g., Yates 1946, 1953; Deming 1950; Cochran 1963; Sukhatme and Sukhatme 1970). Empirical investigations have been restricted to terrestrial systems, particularly to surveys of vegetation types or timber volumes (e.g., Hasel 1938; Osborne 1942; Finney 1948b, 1950; Numata and Nobuhara 1952; Bourdeau 1953; Milne 1959). The results from these studies indicate that randomly located systematic sampling often gives more accurate estimates than other procedures (Hasel 1938; Osborne 1942; Madow 1946; Yates 1946, 1948; Finney 1948a; Bordeau 1953; Milne 1959; Grieg-Smith 1964) especially when the sampled population has positive correlation between neighboring units (Cochran 1946; Milne 1959; Sukhatme and Sukhatme 1970). Because of the greater precision and greater convenience of systematic sampling, some workers have recommended its use for terrestrial surveys (Hasel 1938; Yates 1946; Milne 1959). On the other hand, it has been shown that irregular distributions or pronounced patterns of variation, especially periodicity or linear trends, may cause systematic designs to give biased estimates or estimates of reduced precision (Madow and Madow 1944; Finney 1950; Bourdeau 1953; Sukhatme and Sukhatme 1970); nor does the precision necessarily improve with increasing sample size (Madow 1946; Bordeau 1953).

Of the random designs, SR generally offers greater precision than unrestricted random sampling (Yates 1953; Milne 1959) and, with a constant number of samples, this precision increases as the number of strata increases (Yates 1953). The most precise design is one with one sample per strata, but this (like a systematic sample) offers no internal estimate of error (Finney 1948a, b).

The success of systematic sampling clearly depends upon the nature of the sampled population. If individuals or properties in a population are distributed at random, all strategies will be equivalent. Pronounced pattern, however, may increase or decrease the effectiveness of systematic designs. Thus, quite aside from the theoretical objections to systematic sampling, uninformed application of any systematic sampling is to be discouraged.

Although Strickland (1968) warned that discrete samples may give a poor representation of the vertical distributions of highly stratified substances, such as chlorophyll, a thorough study of the consequences of systematic sampling in the ocean has not been conducted, even though most populations have marked gradients, especially along the vertical axis. This may be attributed to the logistical difficulties of enumerating an oceanic population in its entirety, in contrast to a timber stand in which every individual may be observed, counted, measured, and mapped.

The present study is restricted to the consequences of applying RSS in the vertical direction. The distribution investigated is that of chlorophyll in an oligotrophic oceanic environment. Total chlorophyll in the water column is a frequently used index of plant crop and it is most often estimated from a series of restricted systematic samples. The major question is whether such sampling produces any bias in the estimate of total chlorophyll, or whether the temporal and spatial heterogeneity of the chlorophyll distribution is sufficient to average out the biases of individual determinations. Of secondary concern is whether there is a significant difference in precision or accuracy between estimates derived from RSS and those derived from SR.

The area of study is the North Pacific Central Gyre in the vicinity of lat. 28°N, long. 155°W. The region is one of relatively low spatial and temporal variability (Venrick et al 1973; Gregg et al. 1973; McGowan and Williams 1973; Eppley et al. 1973; Haury 1976). Thus, it is an environment in which any adverse characteristics of RSS are expected to be magnified. The general features of the distribution of chlorophyll in the North Pacific Central Gyre have been summarized (Venrick et al. 1973). Most of the year, surface concentrations are low (0.02-0.06 mg/m^3), and there is a narrow subsurface maximum layer (0.10-0.20 mg/m^3) centered between 90 and 120 m.

The present study was conducted in two parts. In part A, a computer was used to sample nine semiartificial populations derived from continuous vertical profiles of chlorophyll fluorescence. Changes in the fluorescence per extractable chlorophyll unit with depth (Kiefer 1973) and smoothing of small-scale features during the pumping procedure result in a profile which represents only the grosser features of the true distribution. From the vertical profiles, the total population along the vertical axis was calculated, allowing the accuracy of various sampling strategies to be determined directly. Study B was conducted in the field where restricted systematic and stratified random samples were collected simultaneously from the population. In this study, a real population was studied but the total population could only be approximated.

METHODS

Analytical Procedures

Chlorophyll a was determined fluorometrically according to the procedure of Yentsch and Menzel (1963) as modified by Holm-Hansen et al. (1965). Water for discrete, extracted chlorophyll samples was obtained with Nansen bottles. Water for continuous vertical profiles was obtained with the seawater pumping system described by Beers et al. (1967) and was passed through a fluorometer equipped with a flowthrough door.

Study A

The chlorophyll fluorescence profiles were taken during September 1968, on 9 consecutive days during which time the ship followed two drogues which were set at 10 m depth to follow the mixed layer. These were launched at lat. 27°00′N, long. 155°18′W and moved in a northwesterly direction at speeds between 0.5 and 1.5 kn covering 345 km in 9 days. The profiles were not made at the same time of day. The closest two profiles were separated by 13 h, the most distant by 40 h. Additional aspects of these profiles and accessory data have been published (Scripps Institution of Oceanography 1974).

The fluorescence profiles were read at 1-m intervals and translated into units of approximate chlorophyll down to a depth of 180 m. In order to offset the increase in fluorescence per unit of extractable chlorophyll with depth, one conversion equation was used down to and including the chlorophyll maximum and another below the maximum. The conversion factors were determined by analysis of chlorophyll extracted from discrete water samples collected periodically during the cruise. The surface value of each continuous profile was set to 0.03 mg/m^3 and the minimum value below the maximum to 0.01 mg/m^3; these were the mean values of extracted chlorophyll observed at the surface and at 200 m, respectively. The horizontal scale was adjusted to bring the mean maximum value of all profiles to 0.156 mg/m^3, the average maximum of the discrete samples. A typical adjusted profile is presented in Figure 2.

These semi-artificial populations were sampled with four stratified random designs (Table 1). The success of SR depends upon the extent to which the strata can be made internally homogeneous. In an attempt to achieve this, the stratum boundaries of SR-1 and SR-2 were determined as much as possible by the hydrographic, biological, and chemical

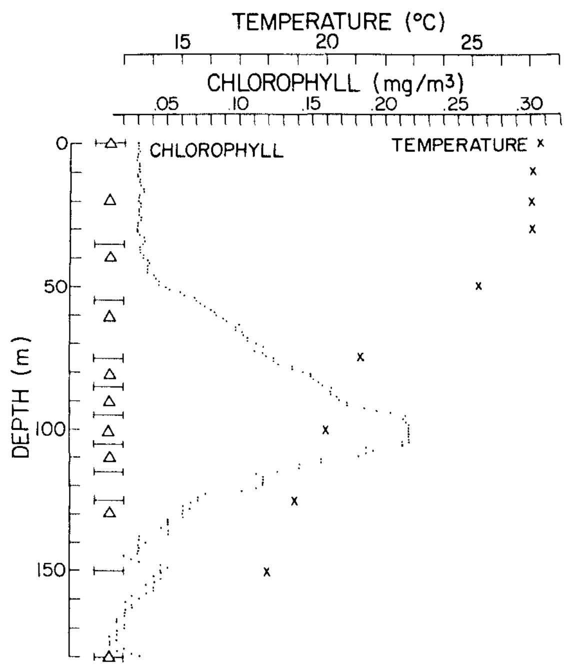

FIGURE 2.—A typical population of chlorophyll values derived from a continuous profile of fluorescence (27 September 1968) and sampled in study A, together with the temperature values from the associated hydrocast. Triangles indicate the location of samples in restricted systematic design 3; bars represent the boundaries of strata used in stratified random design 1.

TABLE 1.—Systematic and stratified random sampling designs used in studies A and B.

Systematic:	RSS-1	RSS-2	RSS-3	RSS-4	Stratified random:	SR-1	SR-2	SR-3	SR-4
	0	0	0	0		0	0	0	0
	20	10	20	45		35	75	15	45
	40	25	40	65		55	95	45	90
Sample	60	35	60	80	Stratum	75	105	75	110
depths	80	50	80	90	boundaries	85	125	90	130
(m)	100	60	90	100	(m)	95	180	100	180
	120	75	100	110		105		110	
	140	100	110	120		115		120	
	160	125	130	137		125		130	
	180	180	180	180		150		150	
						180		180	

characteristics of the environment (Figure 2), and larger strata were assigned to the layers in which environmental gradients were small and several narrow strata were placed in the region of the chlorophyll maximum. The 35-m boundary marked the average depth of the mixed layer; 95 m was the approximate depth of penetration of 1% of the surface radiation, and 125 m represented the beginning of the nutricline. Design SR-1 consisted of 10 strata, each with one sample; in design SR-2, adjacent strata were lumped giving five strata with two samples in each. Designs SR-3 and SR-4 were those used in study B (below) and were thus based on environmental characteristics observed at that time. Each of the nine populations was sampled 20 times with each of the stratified random designs. To facilitate comparison with the systematic samples, for which there was only one cast of each design per profile, it was desirable to examine a series of unreplicated stratified random samples. For this purpose, 10 subsets were selected at random from the replicate casts, each subset containing nine stratified random casts, one from each population. Total chlorophyll was calculated from the mean (arithmetic) concentration per strata times the width of that strata, summed over all strata. This is the classical procedure for summarizing data collected by SR.

Four RSS designs were employed: RSS-1, one sample at the surface and every 20 m thereafter; RSS-2, the design actually employed in September 1968 in which the cast was partially determined by standard hydrographic depths; RSS-3, a design which was based upon complete knowledge of the vertical distributions which were being sampled and which was derived from application of the general rules of sample allocation, i.e., samples were concentrated in the region of maximum variability (the chlorophyll maximum layer); RSS-4, a design based on stratified random design 1 (and therefore more strictly comparable to it) with a sample at the top of the upper stratum (0 m) and bottom of the lowest stratum (180 m) and at the center of all intermediate strata.

Two methods of calculating total chlorophyll from systematic samples were investigated. In the first, the layer between adjacent samples was represented by the arithmetic mean of the two samples (equivalent to integration with linear interpolation). In the second, the layer was represented by the geometric mean of adjacent samples. This latter procedure is sometimes recommended when the population exhibits large, nonlinear changes between adjacent samples. A comparison of the two procedures was made in study A, on the basis of which the method using geometric means was rejected.

Study B

Study B, conducted in June 1977, combined two 10-sample designs, one restricted systematic (RSS-1) and one stratified random (SR-3 or SR-4) into a single 20-bottle cast. The strata boundaries were primarily determined from two preliminary 18-bottle casts which defined the regions of chlorophyll gradients and from a single STD trace which defined hydrographic strata (Figure 3). As in study A, narrower strata were established at the depths of maximum gradients of chlorophyll (the region of the maximum layer). The major differences between designs SR-1 and SR-2 and designs SR-3 and SR-4 were due to a shallower mixed layer and broader, deeper maximum layer observed in June 1977.

Over a period of 21 days, a total of 18 casts were made, 9 employing RSS-1 and SR-3 and 9 employing RSS-1 and SR-4. All casts were located within a rectangle bounded by lat. 28°21.6′N and 28°45.9′N, and by long. 155°14.0′W and 155°33.5′W. Fourteen casts were taken in conjunction with another program between the hours of 2200 and 0300 (with one exception, delayed by winch failure until 0550). Twelve casts were

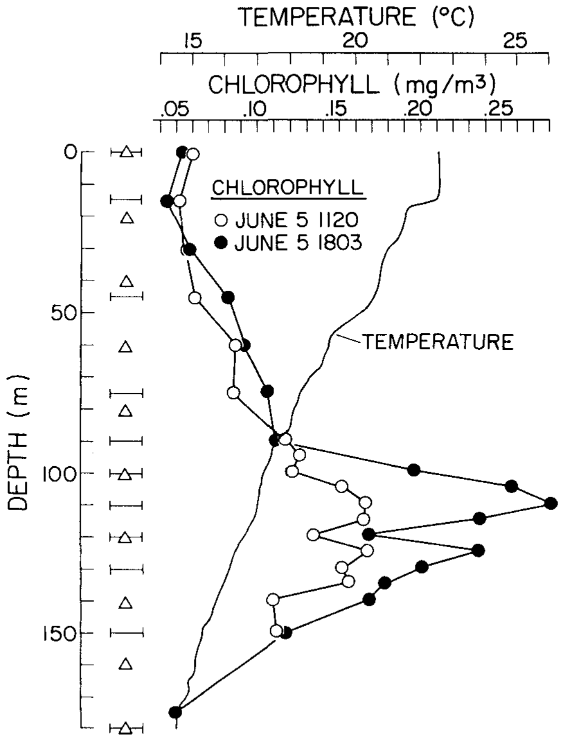

FIGURE 3.—Chlorophyll values observed in two 18-bottle casts preliminary to study B, together with the temperature trace from the associated STD lowering. Triangles indicate location of samples in restricted systematic design 1; bars represent the boundaries of strata used in stratified random design 3.

paired, taken within a few hours and within 3 n.mi. of each other. These have been considered replicate casts.

When the combined systematic-stratified random design called for bottles to be spaced more closely than 3 m, it was necessary to use a messenger heavier than the standard Nansen messenger (such as a Niskin bottle messenger) in order that it develop enough momentum to trip the second bottle. When both sampling designs called for the same depth, the extra bottle was arbitrarily positioned, usually filling in the largest gap in the region of the chlorophyll maximum layer. This "free" sample was used only in the calculations of total chlorophyll in the water column.

Statistical Procedures

Bias is evaluated by the consistency with which n observations (x_i, $i = 1, n$) from a given sampling design fall above or below the true population value and may be measured as a percent of the true value (θ):

$$\left[\frac{\sum (x_i - \theta)}{n} \times 100\%\right] / \theta .$$

Precision is measured by the variance of a series of n observations about their mean ($\bar{x}$):

$$\frac{\sum (x_i - \bar{x})^2}{n-1} .$$

In an analogous way, accuracy is measured by the mean square deviation of a series of observations from the true population total:

$$\frac{\sum (x_i - \theta)^2}{n-1} .$$

Both accuracy and precision are inversely related to their statistical measures, increasing as the numerical value of the measure decreases. Since most scientists are used to thinking in terms of variances and sums of squares, it did not seem desirable to invert these measures to achieve direct correspondence.

In the analysis of the results, limited use was made of the parametric analysis of variance. Most statistical tests were nonparametric tests which make few assumptions about the characteristics of the data (e.g., Dixon and Massey 1957; Tate and Clelland 1957; Conover 1971; Hollander and Wolfe 1973). Unless stated otherwise, the probabilities associated with conclusions in the text are derived from the binomial distribution with $p = ½$.

In several analyses in these studies, the problem of multiple testing arose, as when all four systematic designs were tested for bias. Unfortunately, the tabulation on most nonparametric procedures is not sufficiently complete to allow correction for multiple testing to be made without making the tests extremely conservative. Since this was deemed undesirable, the probabilities given for the statistical tests are uncorrected. It is unlikely that this makes any real difference in the outcome of these studies which gain most of their force from the similarity of results in the two approaches.

RESULTS

Study A

Integration of values

The results of study A are summarized in Table 2. The total chlorophyll values derived from the four systematic sample designs were calculated by integration with linear interpolation (i.e., using the arithmetic mean of adjacent samples to represent the average chlorophyll in the stratum between them). Use of the geometric mean in this calculation resulted in the true total being underestimated 27 out of 36 times ($P = 0.01$). Nor was there any increase in accuracy (the resultant accuracies, based on use of the geometric mean, were 0.538, 0.987, 0.488, and 0.752 for RSS-1 through RSS-4). The use of the geometric mean in the calculation of total chlorophyll does not appear to be justified.

Bias

The biases observed in the eight sampling strategies are summarized in Table 3. Of the four restricted systematic designs, only RSS-2 gave no signs of bias. Design RSS-3, the "best informed" design, overestimated the true population total in eight of the nine trials ($P<0.05$). RSS-1 overestimated the population only five out of nine times, but the overestimates were clustered toward the end of the series and the underestimates toward the beginning. This temporal trend lies just outside the usual level of significance (run test; $P<0.10$) but it indicates that the time period necessary for the population to provide "random" variability of sufficient magnitude to eliminate bias may be of the order of several days or 100-200 km. The magnitudes of the biases were −4.0% for the period 19-21 September and +3.7% for the period 24-27 September. Similarly, the bias introduced by using RSS-3 to estimate the true population total for 19-25 September was −3.6%.

The peculiar periodicity of bias seen in RSS-4 also indicates a nonrandom interaction between the sampling design and the sampled population

TABLE 3.—Bias of systematic and stratified random sampling designs, Study A.

Date (1968)	Systematic designs				Stratified random designs			
	RSS-1	RSS-2	RSS-3	RSS-4	SR-1	SR-2	SR-3	SR-4
19 Sept.	−	−	+	−	+	− −	−	+
20 Sept.	−	−	+	+	−	− +	[1]0	−
21 Sept	−	−	+	−	+	− +	[1]0	+
22 Sept.	+	+	+	+	−	− +	[1]0	+
23 Sept.	−	−	+	−	−	− −	−	−
24 Sept.	+	−	+	+	+	− −	+	+
25 Sept.	+	+	+	[1]0	−	− −	−	+
26 Sept.	+	−	−	−	+	− +	−	+
27 Sept.	+	+	+	+	+	− +	+	−

[1]Estimate ≡ true value.

TABLE 2.—Results of study A, a computerized simulation sampling study. The estimated parameter (θ) is total chlorophyll above 180 m; units are milligrams per square meter; time is local time.

Date (1968)	Time	True value (θ)	Systematic designs One cast each				Stratified random designs Means and variances (in parentheses) of 20 replicates			
			RSS-1	RSS-2	RSS-3	RSS-4	SR-1	SR-2	SR-3	SR-4
19 Sept.	1719	8.50	8.00	8.00	8.62	8.39	8.52 (0.345)	8.38 (2.460)	8.39 (0.642)	8.59 (0.050)
20 Sept.	2312	10.31	9.91	9.98	11.12	10.71	10.29 (0.244)	10.18 (0.987)	10.31 (0.182)	10.24 (0.513)
21 Sept.	2335	7.35	7.21	6.69	7.57	7.33	7.36 (0.059)	7.16 (0.950)	7.35 (0.215)	7.36 (0.466)
22 Sept.	2351	6.89	7.23	7.72	7.45	7.18	6.85 (0.072)	6.88 (0.151)	6.89 (0.080)	6.92 (0.129)
23 Sept.	2025	8.83	7.97	7.51	9.03	8.47	8.62 (0.801)	8.68 (1.780)	8.82 (0.341)	8.78 (1.227)
24 Sept.	0900	9.68	9.70	9.20	9.94	10.56	9.81 (0.178)	9.57 (1.490)	9.85 (0.406)	9.79 (1.841)
25 Sept.	0800	11.00	11.84	12.08	11.08	11.00	10.86 (0.263)	10.90 (1.436)	10.92 (0.443)	11.16 (0.687)
26 Sept.	0830	13.85	14.23	13.14	13.36	13.06	13.92 (1.123)	13.23 (5.323)	13.65 (2.289)	13.90 (3.954)
27 Sept.	2400	13.90	14.47	14.56	14.94	14.78	14.17 (0.490)	13.60 (3.676)	14.06 (1.409)	13.83 (3.348)
Accuracy $\frac{\sum(x_i - \theta)^2}{n-1}$			0.308	0.695	0.308	0.320	[1]0.574	[1]1.464	[1]0.779	[1]1.662
Precision $\frac{\sum(x_i - \bar{x})^2}{n-1}$		6.441	8.113	7.729	6.537	6.725	[1]8.004	[1]6.312	[1]7.036	[1]8.540

[1]Mean values from 10 sets of unreplicated casts.

(run test, $P = 0.10$). With the sampling interval employed here, the biases of individual estimates average out over the entire study. Had the interval been twice as large, a consistent overestimate or underestimate would have resulted, with respective magnitudes of +5.8% and −1.9%, until 25 September when the phase relationship appears to have shifted.

Tables 2 and 3 also present the results of the four stratified random designs, based upon the means of 20 replicates. The consistent underestimates resulting from SR-2 were sufficiently unexpected that a second series of 20 SR-2 samples were drawn from each population. This series showed no evidence of bias and, thus, it appears that the initial results were the product of random chance.

Precision

Precision, in its strictest sense, could only be examined in the case of the stratified random designs, for which replicates were available. The designs employing 10 strata, each with one sample, SR-1 and SR-3, offered greater precision than designs with fewer strata. However, there was a highly significant concordance (Kendell coefficient, $P<0.01$) between the precisions of all designs with respect to the profiles giving the most precise result. Examination of the individual profiles indicated that the precision of the results was inversely related to the strength of the chlorophyll maximum and to the amount of small-scale variability along the vertical axis, or, in other words, to the structural complexity of the population. Later, the accuracy of the systematic designs (discussed below) was found to show the same relationship.

For all stratified random designs, the variance between replicates was trivial compared with the variance between the nine populations. Analyses of variance gave $f_{8,19}$ ratios ranging from 54 to 344 (all $P<<0.01$). When all nine profiles were considered to be replicates of the same population, the variance between the nine estimates from each systematic cast could be compared with the variance between single stratified random casts, one from each population (Figure 4A). On this scale, there were no differences in precision between any of the sampling designs. The large variation between populations masked any difference in performance. Thus, when the concept of the sampled population is expanded to include spatial and temporal variations, RSS appears to offer neither advantages nor disadvantages with respect to precision of estimates.

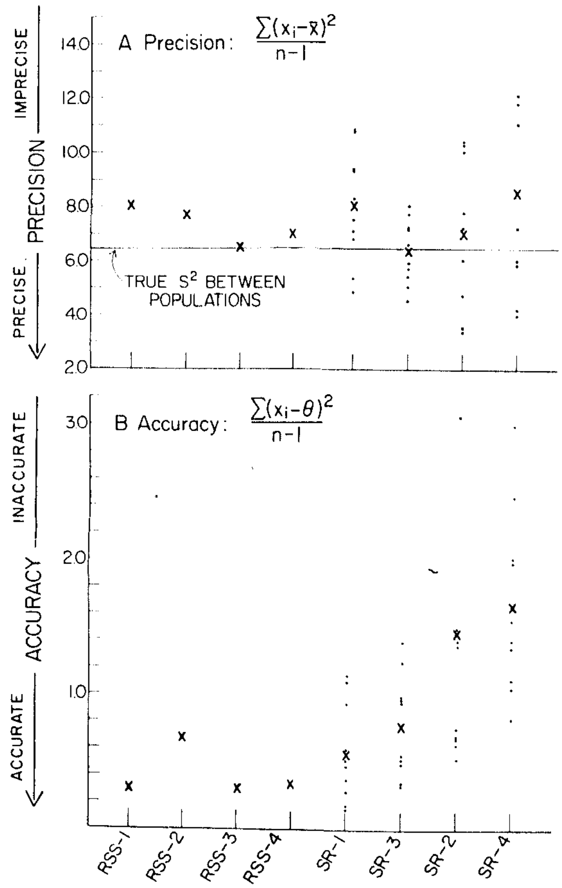

FIGURE 4.—The results of the computer simulation sampling study, study A, showing the relative precisions and accuracies of the four restricted systematic sampling designs (RSS) and four stratified random designs (SR).

Accuracy

The accuracy of the various designs was also compared using sets of unreplicated stratified random casts (Figure 4B). The greater accuracies of stratified random designs SR-1 and SR-3 relative to SR-2 and SR-4 undoubtedly reflected their greater precision; and perhaps the greater accuracy of SR-1 relative to SR-3 was due to selection of more appropriate strata. The systematic designs were generally more accurate than the stratified random designs. Only stratified random design SR-1 achieved the accuracy of the systematic designs.

Most of the chlorophyll work in the central Pacific has been based upon 12 or more sampled depths. Thus, it was encouraging to find that as few as 10 depths, regardless of the sampling strategy, gave a generally satisfactory picture of the amount of chlorophyll in the water column. Of nearly 400 estimates from individual casts, 76% fell within ±10% of the true value. This percent increased to 85% for stratified designs SR-1 and SR-3 and to 94% for the 36 systematic casts. However, to the extent that these fluorescence profiles underestimate the structural complexity of the true chlorophyll distribution, these results probably overestimate the accuracies of the designs.

Study B

The results of the field study were remarkably similar to those of the computer study (Table 4). Bias and accuracy were investigated by assuming that the entire population was exactly represented by the 20 samples in one cast (systematic samples plus stratified random samples plus "free" samples). The results of study A indicate that the discrepancy is not likely to be severe.

Bias

When the 18 casts are considered in chronological sequence, it is evident that RSS tended to deviate from the true value in the same direction on adjacent casts. The direction of bias was the same within five of the six pairs of replicate casts ($0.05<P<0.10$) and a run test over the entire sequence was significant ($P = 0.05$). The absolute magnitude of the bias which would result were the five replicate pairs considered estimates of five population totals ranged from 0.3% to 3.7% with a mean of 2.0%. Within the restricted spatial area of this survey, between 2 and 8 days appear necessary for the natural fluctuations of the population to be sufficient to average out bias inherent in RSS.

Precision

Precision was investigated by means of the six pairs of replicate casts. Stratified random design SR-3 with one sample per strata was more precise than the design with two samples per strata ($f_{3,3} = 6.5$, $P<0.10$). The precision of the systematic design was intermediate and was not significantly different from either.

TABLE 4.—Results of study B, a field sampling study. The estimated parameter is total chlorophyll above 180 m and the true value (θ) is estimated from the 20 combined samples of the two designs; units are milligrams per cubic meter.

Date (1977)	Local time	$\hat{\theta}$	x_i bias RSS-1	SR-3	SR-4
5 June]	2345	18.87	18.73−	18.25−	
6 June]	0220	17.68	17.10−	18.69+	
8 June	0033	15.54	14.06−	18.21+	
9 June	0236	18.11	17.23−	17.55−	
9 June]	2241	16.70	16.22−	17.19+	
10 June]	0550	15.47	14.75−	16.68+	
13 June]	2208	13.26	13.29+	12.31−	
14 June]	0035	13.42	13.47+	13.33−	
15 June	2203	14.53	15.91+	11.19−	
19 June]	2149	11.50	10.91−		9.03−
20 June]	0100	10.29	10.78+		10.78+
21 June	2246	14.25	14.73+		11.42−
22 June	1107	10.67	11.00+		10.17−
23 June]	2333	13.60	13.41−		12.97−
24 June]	0133	12.52	12.16−		12.09−
24 June]	1505	16.83	17.11+		18.11+
24 June]	1632	16.98	17.69+		15.72−
26 June	0822	13.32	13.58+		13.10−
Accuracy:	$\frac{\sum(x_i - \theta)^2}{n-1}$		0.45	2.83	2.31
Precision [MS w/in pairs]:	(6/5-6/15)	0.492	0.808	0.249	
	(6/19-6/25)	0.442	0.319		1.591

] indicates pair of replicate casts.

Accuracy

The accuracies of the two stratified random designs, as measured against the total chlorophyll estimated from all 20 samples, were similar, but the systematic design RSS-1 was significantly more accurate than either (signed rank test, $P<0.05$). Possibly the greater accuracy of the systematic designs, seen in both studies, might be partially attributable to the different arithmetic formulae with which total chlorophyll was calculated since these give somewhat different weights to the individual samples. To test this, the estimates from both the restricted systematic and the stratified random designs were calculated by linear integration. This did not alter the relative accuracies of RSS-1 and SR-3 in study B or of RSS-1 and SR-1 in study A, nor did it eliminate the bias apparent in RSS-1. Thus, it appears to be the sample location rather than the formula by which total chlorophyll is calculated that is responsible for the greater accuracies of the systematic designs.

DISCUSSION

These studies indicate that there is a potential for biased estimates to be derived from systematic samples collected from a planktonic ecosystem.

While it is recognized that the results suffer from small sample sizes, they gain considerable force from the fact that two different approaches give quite similar conclusions.

It appears that RSS, when applied to the vertical distribution of chlorophyll, may actually give estimates of total chlorophyll which are, on the average, closer to the true value than are estimates based on SR. This is consistent with results from terrestrial systems. Stratified random designs frequently result in pairs of samples falling closely adjacent to one another. In populations which are varying continuously, the information contained in such an adjacent pair is largely repetitious. Such redundancy is avoided in RSS because the spacing between adjacent samples is controlled. The relative performance of SR is expected to improve in populations with more discontinuous distributions such that the strata may be defined to be internally homogeneous, giving maximum precision. On the other hand, and perhaps more important, there appears to be a potential for bias in systematic designs, especially when the sampling occurs within restricted spatial or temporal intervals. In the central Pacific, which is relatively homogeneous in time and space, the bias was detectable, but the magnitude was small. Unfortunately, the results cannot be generalized to other environments. The bias may be expected to diminish as increasing environmental complexity increases the small-scale variability of the sampled population. Whether or not bias is therefore negligible in more complicated neritic environments remains to be investigated.

On the other hand, in planktonic environments the natural fluctuations produce variability between replicate casts which is generally large relative to the experimental error associated with a single cast. The increased accuracy of systematic designs is not likely to result in a detectable increase in the precision of the estimate of the mean of several samples. Depending upon the goals of a study, it may also be true that the magnitude of the bias introduced by systematic sampling is insignificant. For instance, it appears that RSS as routinely used on large-scale oceanic surveys probably introduces no serious error. However, increasing attention is being focused on small-scale, local phenomena, and the routine use of RSS for these studies deserves examination. Possible effects of the interaction between bias and sampling scale found in this study include overestimation of fluctuations in total population (if a positive bias occurs with higher populations and a negative bias with lower populations), underestimation of fluctuations (if a positive bias occurs with lower populations and a negative bias with higher populations), or production of artificial fluctuations when in fact the population is stable. Whether or not such artificial effects of RSS might be important enough to overshadow the gain afforded in terms of ease of sample location and data analysis depends upon the particular study under consideration.

Study A demonstrated the dependence of the success of a sampling design upon the interaction of that design with the structure of the population being sampled; thus, it would seem that intelligent application of knowledge about the sampled population should improve the design. It was, therefore, disconcerting to find that RSS every 20 m, RSS-1, consistently performed as successfully as did RSS-3 which was designed by a presumably experienced worker (the author) with total knowledge about the population to be sampled. We must conclude either that the location of samples in RSS-3 was not as intelligent as it might have been, or that the natural variability of the population makes intelligent placement of systematic samples impossible. (The latter interpretation has a certain appeal.) In contrast, the dependence of SR on the selection of strata is apparent in both studies. Many narrow strata give more precise and accurate estimates than do fewer, larger ones, undoubtedly because, in the presence of strong vertical gradients, they better satisfy the criterion of internal homogeneity. The most precise and accurate results are obtained when the number of strata is equal to the number of samples. The disadvantage of this strategy is the absence of direct estimates of variability within strata from which to calculate confidence intervals around the final estimate. This is not usually of concern in planktonic work because small-scale patchiness is of such magnitude that more useful estimates of precision are obtained directly from replicate casts, and thus apply to some spatial and temporal interval, rather than to a single cast.

The logistics of SR in the open ocean presented few major difficulties. The task of allotting samples randomly to strata was time consuming. After the first few casts, the job of sample design was relegated to the computer. Preparation of the cast card demanded more than routine caution (although minor errors were readily assimilated into the randomization procedure). The use of random sampling also precluded routine sharing of water samples with others whose programs were designed around standard depths, and, with the present high cost of ship time, the pressure for sharing samples is often considerable. Indeed, this study was possible primarily because there was an unusual amount of excess wire time available.

On the other hand, there were advantages to SR quite apart from the general merriment caused by the unorthodox bottle spacing. The results from the occasional closely spaced samples gave continual insight into the vagueries of small-scale vertical stratification. For instance, within the top 30 m, samples separated by 1- and 2-m intervals differed by <3% of their mean. In the region of the chlorophyll maximum, between 75 and 125 m, the same intervals produced deviations of 30% and 40%, indicating sharp layers and frequent inversions.

It is not within the scope of this paper to make generalizations concerning the use of systematic sampling versus SR. The potential advantages and disadvantages of each have been demonstrated in one environment. It is the responsibility of all researchers to evaluate the use of RSS in reference to their specific programs. If a potential problem is recognized, it is hoped that it will stimulate a preliminary sampling study to examine directly the consequences of RSS in the biological system of interest. This worker's experience with SR was satisfactory, and the potential for biased data is sufficiently undesirable that effort will be made to incorporate randomization into future sampling designs. It is hoped that the untidy profiles which will result will be accepted with understanding by the scientific community.

ACKNOWLEDGMENTS

This work was supported in part by the Office of Naval Research Grant N00014-75-c-0152, and in part by the Marine Life Research Program, the Scripps Institution of Oceanography's part of the California Cooperative Oceanic Fisheries Investigations, which are sponsored by the Marine Research Committee of the State of California, and by the National Science Foundation Grant GB-2861.

Michael M. Mullin and Thomas Hayward critically reviewed the manuscript. Special thanks are due to Hayward for unfaltering cheer and enthusiasm and occasional well-needed prompting during 0300 chlorophyll casts.

LITERATURE CITED

BEERS, J. R., G. L. STEWART, AND J. D. H. STRICKLAND.
1967. A pumping system for sampling small plankton. J. Fish. Res. Board Can. 24:1811-1818.

BOURDEAU, P. F.
1953. A test of random versus systematic ecological sampling. Ecology 34:499-512.

COCHRAN, W. G.
1946. Relative accuracy of systematic and stratified random samples for a certain class of populations. Ann. Math. Stat. 17:164-177.
1963. Sampling techniques. 2d ed. John Wiley and Sons, Inc., N.Y., 413 p.

CONOVER, W. J.
1971. Practical nonparametric statistics. John Wiley and Sons, N.Y., 462 p.

DEMING, W. E.
1950. Some theory of sampling. John Wiley and Sons, Inc., N.Y., 602 p.

DIXON, W. J., AND F. J. MASSEY, JR.
1957. Introduction to statistical analysis. 2d ed. McGraw-Hill, N.Y., 488 p.

EPPLEY, R. W., E. H. RENGER, E. L. VENRICK, AND M. M. MULLIN.
1973. A study of plankton dynamics and nutrient cycling in the central gyre of the North Pacific Ocean. Limnol. Oceanogr. 18:534-551.

FINNEY, D. J.
1948a. Volume estimation of standing timber by sampling. Forestry 21:179-203.
1948b. Random and systematic sampling in timber surveys. Forestry 22:64-99.
1950. An example of periodic variation in forest sampling. Forestry 23:96-111.

GREGG, M. C., C. S. COX, AND P. W. HACKER.
1973. Vertical microstructure measurements in the Central North Pacific. J. Phys. Oceanogr. 3:458-469.

GREIG-SMITH, P.
1964. Quantitative plant ecology. 2d ed. Butterworth & Co., Lond., 256 p.

HASEL, A. A.
1938. Sampling error in timber surveys. J. Agric. Res. 57:713-736.

HAURY, L. R.
1976. A comparison of zooplankton patterns in the California Current and North Pacific Central Gyre. Mar. Biol. (Berl.) 37:159-167.

HOLLANDER, M., AND D. A. WOLFE.
1973. Nonparametric statistical methods. John Wiley and Sons, N.Y., 503 p.

HOLM-HANSEN, O., C. J. LORENZEN, R. W. HOLMES, AND J. D. H. STRICKLAND.
1965. Fluorometric determination of chlorophyll. J. Cons. 30:3-15.

KIEFER, D. A.
1973. Fluorescence properties of natural phytoplankton populations. Mar. Biol. (Berl.) 22:263-269.

MADOW, L. H.
1946. Systematic sampling and its relation to other sampling designs. J. Am. Stat. Assoc. 41:204-217.

MADOW, W. G., AND L. H. MADOW.
1944. On the theory of systematic sampling, I. Ann. Math. Stat. 15:1-24.

MCGOWAN, J. A., AND P. M. WILLIAMS.
1973. Oceanic habitat differences in the North Pacific. J. Exp. Mar. Biol. Ecol. 12:187-217.

MILNE, A.
1959. The centric systematic area-sample treated as a random sample. Biometrics 15:270-297.

NUMATA, M., AND H. NOBUHARA.
1952. Studies on the coastal vegetation at Nyijigahama (Report 1). Bot. Mag. Tokyo 65:149-157.

OSBORNE, J. G.
1942. Sampling errors of systematic and random surveys of cover-type areas. J. Am. Stat. Assoc. 37:256-264.

SCRIPPS INSTITUTION OF OCEANOGRAPHY.
1974. Physical, chemical and biological data, Climax I Expedition, 19 September-28 September 1968. SIO Ref. 74-20, 41 p.

STRICKLAND, J. D. H.
1968. A comparison of profiles of nutrient and chlorophyll concentrations taken from discrete depths and by continuous recording. Limnol. Oceanogr. 13:388-391.

SUKHATME, P. V., AND B. V. SUKHATME.
1970. Sampling theory of surveys with applications. 2d ed. Iowa State Univ. Press, Ames, 452 p.

TATE, M. W., AND R. C. CLELLAND.
1957. Nonparametric and shortcut statistics in the social, biological, and medical sciences. Interstate Printers and Publishers, Inc., Danville, Ill., 171 p.

VENRICK, E. L., J. A. MCGOWAN, AND A. W. MANTYLA.
1973. Deep maxima of photosynthetic chlorophyll in the Pacific Ocean. Fish. Bull., U.S. 71:41-52.

YATES, F.
1946. A review of recent statistical developments in sampling and sampling surveys. J. R. Stat. Soc., Lond. 109:12-43.
1948. Systematic sampling. Philos. Trans. R. Soc. Lond., Ser. A, 241:345-377.
1953. Sampling methods for censuses and surveys. 2d ed. Hafner Publishing Co., N.Y., 401 p.

YENTSCH, C. S., AND D. W. MENZEL.
1963. A method for the determination of phytoplankton chlorophyll and phaeophytin by fluorescence. Deep-Sea Res. 10:221-231.