UCLA
UCLA Previously Published Works

Title
Sampling Procedures for Regional Surveys: A Problem of Representativeness and Effectiveness

Permalink
https://escholarship.org/uc/item/93w9g3v4

Journal
Journal of Field Archaeology, 13(4)

ISSN
0093-4690

Author
Read, Dwight W

Publication Date
1986

DOI
10.2307/530171

Peer reviewed

Sampling Procedures for Regional Surveys: a Problem of Representativeness and Effectiveness

Dwight W. Read
University of California, Los Angeles

In this paper several considerations are discussed that developed out of the sampling program for the Chevelon Archaeological Research Project, located in northern Arizona, aimed at making regional surveys more effective in their recovery of data on site locations. It is argued that a sampling program based on the survey of quadrats should be evaluated in terms of its ability to effectively discover sites, consistent with creating a sample of sites representative of the population of sites in the region.

Effectiveness can be measured against the expected number of sites to be found through the sampling program. Some of the factors having a substantial impact on effectiveness are discussed and it is shown, using the data from the Chevelon Project, that a well-designed sampling program can increase the effectiveness of the survey several-fold over a simple random selection of quadrats. The impact of large areas of low site density on sampling effectiveness is considered in detail and tables are developed for evaluating the cost (measured in terms of the minimum number of quadrats needed) of surveying areas with a low density of sites and for testing whether a stratum having such a low density of sites should be excluded from a probabilistic sampling program. A sampling program aimed at delineating the spatial distribution of sites is outlined as a means to resolve the problem of surveying regions containing large areas, each with low site densities.

Introduction

This paper addresses several considerations aimed at making regional surveys more effective in their recovery of data on site locations. The discussion is based on the sampling program developed by the author for the Chevelon Archaeological Research Project[1] and involves the synthesis of two complementary viewpoints: the mathematical basis for sampling within a statistical, inferential framework; and the archaeological research interests toward which the sampling methodology is aimed.

The mathematical and the archaeological approaches to sampling design are complementary in the sense that the research problem defines what data are needed and the sampling methods provide the means by which these data may be obtained in a manner statistically representative of the total variety of sites and site localities within the region. Statistical representativeness is necessary for parameter estimation, but the latter is needed only when a portion of the whole population is measured. The role of sampling in statistical methodology is to permit inference of the whole from the part when information on the whole cannot be obtained. With a complete survey of all the sites in the region, one obtains the benefits of statistical analysis[2] by directly measuring population parameters without entering into estimation procedures and statistical testing, thereby avoiding the complication of designing an effective probabilistic sampling program for spatially-patterned data.

While complete data are the ideal, they are costly to procure as measured by time and money (e.g., the $50 million estimate for complete mitigation of the approximately 53.1 sq km of Horseshoe Reservoir in Arizona).[3] Or, the regional area may be unmanageably large for a

1. F. Plog, J. Hill, and D. Read, eds., *Chevelon Archaeological Research Project 1971–1972. UCLA Archaeological Survey Monograph* 2 (UCLA Archaeological Survey: Los Angeles 1976).

2. D. Read, "Some Comments on the Use of Mathematical Models in Anthropology," *AmAnth* 39 (1974) 3–15; G. A. Clark, "Quantifying Archaeological Research," in M. Schiffer, ed., *Advances in Archaeological Method and Theory* 5 (Academic Press: New York 1982) 217–273.

3. S. Fuller, A. Rogge, and L. Gregonis, *Orme Alternatives: the Archaeological Resources of Roosevelt Lake and Horseshoe Reservoir. Arizona State Museum Archaeological Series* 98 (Arizona State University: Tempe 1976).
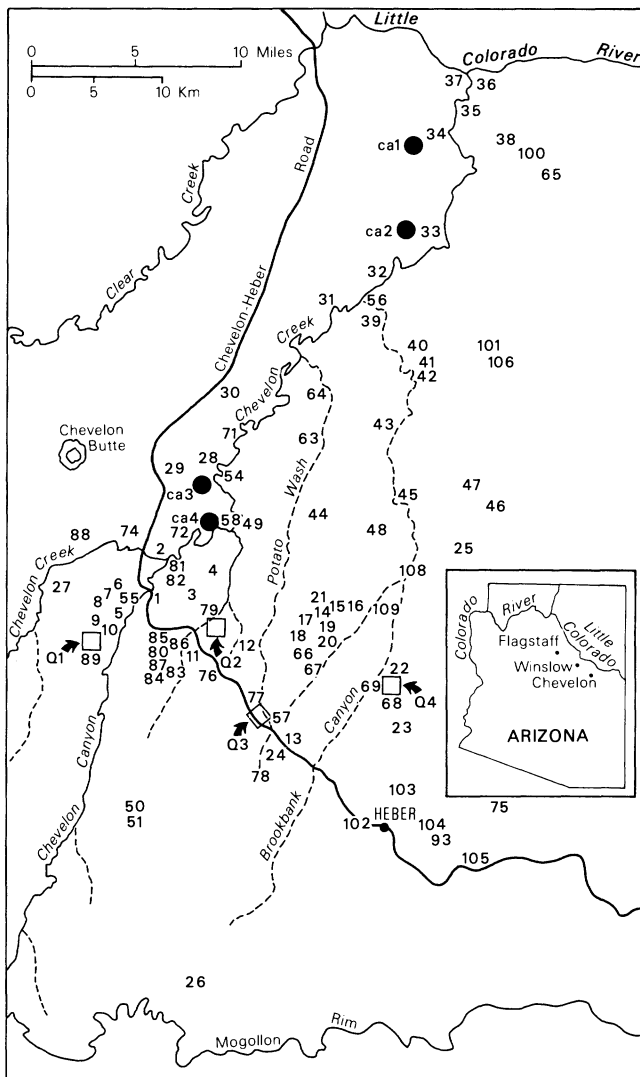
Figure 1. Location of Chevelon drainage and survey units.

Two major guiding principles for the construction of a sampling program for the regional survey can be inferred from these discussions: 1) the sampling procedure should produce a sample of sites statistically representative of the population of sites in the region and 2) the sampling procedure should be efficient in its use of time and other resources. While there is agreement on these principles, their implications are less clear for the construction of a regional sampling program. Though representativeness can be easily maintained using the sample survey method,[6] sampling efficiency for a particular sample design varies markedly according to the spatial distribution of sites in the region, making it difficult to determine the best sampling strategy in the absence of information on the spatial patterning of sites. Nonetheless, general factors especially relevant to designing an effective archaeological sampling program can be identified and their impact on the effectiveness of the sampling program assessed, as will be done in this paper.

The sampling program for the Chevelon Archaeological Research Project (CARP) will be considered in some detail in order to exemplify the archaeological considerations that arise and need to be incorporated when

complete survey (e.g., the Bureau of Land Management survey of archaeological and historical resources for planning purposes in desert areas includes about one-fourth of the state of California).[4] Hence it is often necessary to sample, and a number of suggestions have been advanced about how regional surveys should be done[5] so as to obtain systematic and representative data at the regional level.

4. *The California Desert Conservation Area: Plan Alternatives and Environmental Impact Statement (Draft)* (Bureau of Land Management: Sacramento 1980).

5. L. Binford, "A Consideration of Archaeological Research Design," *AmAnt* 29 (1964) 425–441; C. Redman, "Multistage Fieldwork and Analytical Techniques," *AmAnt* 38 (1973) 61–79; J. Mueller, *The Use of Sampling in Archaeological Survey. SAA Mem* 28 (Society for American Archaeology: Washington, D.C. 1974); J. Mueller, "Ar-

chaeological Research as Clustering Sampling," in J. Mueller, ed., *Sampling in Archaeology* (The University of Arizona Press: Tucson 1975) 33–41; R. Chenhall, "A Rationale for Archaeological Sampling," in J. Mueller, ed., op. cit. (in this note) 3–25; W. Judge, J. Ebert, and R. Hitchcock, "Sampling in Regional Archaeological Survey," in J. Mueller, ed., op. cit. (in this note) 82–183; R. Matson and D. Lipe, "Regional Sampling: a Case Study of Cedar Mesa, Utah," in J. Mueller, ed., op. cit. (in this note) 124–143; D. Read, "Regional Sampling," in J. Mueller, ed., op. cit. (in this note) 45–60; idem, "Review of 'Orme Alternatives: the Archaeological Resources of Roosevelt Lake and Horseshoe Reservoir,'" in *Orme Alternatives: the Archaeological Resources of Roosevelt Lake and Horseshoe Reservoir. Arizona State Museum Archaeological Series* 98 (Arizona State Museum: Tucson 1976) 281–284; M. Schiffer, A. Sullivan, and T. Klinger, "The Design of Archaeological Surveys," *WA* 10 (1978) 1–28; S. Plog, "Sampling in Archaeological Surveys: a Critique," *AmAnt* 43 (1978) 280–285; S. Plog, F. Plog, and W. Wait, "Decision Making in Modern Surveys," in M. Schiffer, ed., *Advances in Archaeological Method and Theory* 1 (Academic Press: New York 1978) 383–421; J. Nance, "Statistical Fact and Archaeological Faith: Two Models in Small Site Sampling," *JFA* 8 (1981) 151–165; idem, "Regional Sampling in Archaeological Survey," in M. Schiffer, ed., *Advances in Archaeological Method and Theory* 6 (Academic Press: New York 1983) 289–356.

6. Basic texts are M. H. Hansen, W. N. Hurwitz, and W. G. Madow, *Sample Survey Methods and Theory* (John Wiley and Sons: New York 1953); L. Kish, *Sampling Survey Methods* (John Wiley and Sons: New York 1965); P. V. Sukhatme and B. V. Sukhatme, *Sampling Theory of Surveys with Applications* (Food and Agricultural Organization: Rome 1970); W. Cochran, *Sampling Techniques* (John Wiley and Sons: New York 1977).

constructing a sampling program (FIG. 1). Some of the issues that arose with the implementation of the CARP sampling program will be used to illustrate factors that have substantial influence on the effectiveness, and hence the design, of a sampling program.

Of these, one finds that information on the nature of the spatial distribution of sites is perhaps the most critical, yet is often a category of information about which little is known in advance, and represents a kind of data difficult to recover from a sampling program oriented toward parameter estimation. A resolution—developed out of our experiences with the CARP sampling program—to the double problem of obtaining the information one needs for designing an effective regional survey sampling program, and the sampling design required to determine spatial patterning, will also be discussed.

Research Goals of the Chevelon Archaeological Research Project

The general goal of the CARP project was to obtain data on the full range of archaeological sites in the Chevelon drainage. These data were (and are) to be used for a variety of research interests centered on viewing the occupation of the Chevelon drainage as a topographically, if not culturally and socially, bounded system. In addition to the main research questions, data from the survey have permitted reexamining what was then conventional wisdom about site variety and diversity, measured spatially and through time, for this part of the Southwest.

Conventional wisdom was based on extensive work with large sites supplemented by limited regional data. By taking the entire Chevelon drainage as the region, with its area of approximately 10,000 sq km (FIG. 1), it was possible to determine if previous work in this area had been biased by choice of a data base which was not inclusive of the full variety of sites and their locations in the area. The probabilistic survey of the Chevelon drainage has clearly shown the implicit bias: "We have learned . . . that typical sites during most time periods had 3 rooms, not the 30 indicated in many syntheses."[7]

As for the research orientation, the goal has been to develop models that effectively account for the space/time/form variation in settlements and settlement patterns found in the drainage. The specific hypotheses that have guided the research have been presented elsewhere[8] and are of a like nature with the general goals and aims

of regional studies that have been discussed extensively in the literature.[9]

From the perspective of the present moment, neither the goals of the project nor the means used to implement them would now be seen as strikingly innovative, yet at the time the project was begun (summer of 1971), surveys of regions on the order of 10,000 sq km were unusual, particularly surveys in which the selection of areas to be investigated was not made on the basis of the expertise of the archaeologist regarding likely site location, but rather through the use of a table of random numbers.

The rationale for the division of a region into small spatial units and then selecting a sample of spatial units for survey on the basis of a table of random numbers has been discussed in detail elsewhere[10] and need not be repeated here other than to note that it is fundamental for assuring that the basic assumption of statistical inferential methods, namely the independence of data observations, is satisfied. Instead, some of the problems will be identified that arose when attempting to incorporate the archaeological research goals within the framework of statistical sampling procedures, and vice-versa.

Sampling Goals of the CARP Project

The general problem addressed by the sampling scheme for the CARP project was to obtain a substantial, representative sample of sites from the Chevelon drainage to be used for a variety of research purposes. At the time the initial survey began little was known about the form and distribution of sites in the drainage. Since the goals of the project were broad and centered on processes affecting site location and site form, the survey of the region could be limited neither by prior considerations about types of sites that might be found, nor by factors thought to constrain site location (but which had never been adequately tested). Hypotheses that were to be addressed by the data from the survey were not couched in terms of specific expectations about site types and their location, but about processes said to structure the

7. F. Plog, "Is a Little Philosophy (Science?) a Dangerous Thing?" in C. Renfrew, M. Rowlands, and B. Segraves, eds., *Theory and Explanation in Archaeology: the Southampton Conference* (Academic Press: New York 1982) 25–34.

8. Plog, Hill, and Read, eds., op. cit. (in note 1).

9. J. Parsons, "Archaeological Settlement Patterns," *AnnRevAnth* 1 (1972) 127–150; M. Webb, "Exchange Networks in Prehistory," *AnnRevAnth* 3 (1974) 357–383; F. Plog, "Systems Theory in Archaeological Research," *AnnRevArch* 4 (1975) 207–224; I. Hodder and C. Orton, *Spatial Analysis in Archaeology* (Cambridge University Press: Cambridge 1976); C. Smith, ed., *Regional Analysis* (Academic Press: New York 1976); G. Johnson, "Aspects of Regional Analysis in Archaeology," *AnnRevAnth* 6 (1977) 479–508; C. Redman, "Archaeological Survey and the Study of Mesopotamian Systems," *JFA* 9 (1982) 375–382.

10. A basic text is J. Mueller, ed., *Sampling in Archaeology* (The University of Arizona Press: Tucson 1975).

form and location of sites. The form (area, architectural features, artifact content, etc.) and location (longitude and latitude, elevation, landform, ecological zone, etc.) of sites were taken as unknowns and the sampling was to provide data on these aspects of the archaeological material in an unbiased and representative manner.

Two complementary questions arise at this point: first, how one samples in a manner that satisfies the demands of statistical inference and, second, how one samples in a manner that satisfies the archaeological requirements. The answer to the former has a rationale given via sampling theory. Simply put, there is one primary criterion that needs to be satisfied from the viewpoint of statistical inference: each datum in the sample is to be obtained in a manner independent of all other sample data. The criterion of independence of data points is necessary for the construction of confidence intervals for parameter estimation and correct assignment of probabilities of Type I errors (rejection of a true null hypothesis) and Type II errors (acceptance of a false null hypothesis). The former probability is also known as the significance level of a statistical test and the latter is related to the power of the statistical test via the relationship,

$$\text{power} = 1 - \text{Prob [Type II error].}[11]$$

In the context of a regional survey such as the CARP survey, independence of data would mean that each sampling unit such as the site (or each set of sites, keeping fixed the number of sites in a set) must be equally likely to be selected. Simple random sampling through enumeration of the units of the population to be sampled, and then selecting units to be included in the sample by means of a table of random numbers, is the basic means for assuring satisfaction of this criterion.

From the archaeological point of view, however, simple random sampling of sites as units is not possible since the population in question, namely the collection of sites in the Chevelon drainage, is unknown in advance. Further, by the nature of the project, information was wanted not only on the sites, but also on their topographic, ecological, temporal, and social contexts. The topographic and ecological contexts are determined from knowing site spatial location and having detailed information on the topographic and ecological aspects of the region (which is part of the data collected during the survey). The social context is inferred from the temporal position, the form, the content, and the spatial

relation of sites to one another. Hence from the archaeological viewpoint the sampling has several goals, only one of which, namely the estimation of parameters such as the number of sites in the region, is directly subsumed under statistical methodology. These various goals need not each lead to one and the same "best" sampling strategy.

The general strategy used in the CARP sampling was to stratify the region according to ecological (e.g., grassland, juniper-piñon, ponderosa pine, and riverine) and topographic (e.g., river bottom and upland) zones, then to divide the strata into rectangular quadrats (100 m × 1000 m in size) and finally to select a random sample of these quadrats to be surveyed.[12] Dividing the region into quadrats and then selecting a random sample of quadrats serves a dual role. First, it is a means by which to sample sites in an unbiased fashion and, second, it sees that measures related to the local topographic and ecological features associated with the site or sites within the quadrat may be taken.

Thus two samples are constructed: 1) a sample of quadrats, composed of a simple random sample of quadrats drawn from the population of quadrats in a stratum; and 2) a sample of sites, not generally comprising a simple random sample from the population of sites in the region, formed through the survey of the sample quadrats.

Sampling of Sites Versus Site Sample Formation

A distinction is made here between the sampling of sites and the formation of a site sample, as two different concepts are involved. By the sampling of sites, what is meant is the enumeration of all sites in the region followed by selection of a sample of these sites according to a sampling scheme, such as simple random sampling, in which the sampling units are the individual sites. In this type of sampling there is agreement between the unit of measurement and analysis and the unit for sampling which is, in both cases, the site. For example, one might want to test the null hypothesis (H_0) that the average site area is independent of ecological zone (e.g., grassland, juniper-piñon, ponderosa pine) versus the alternative hypothesis (H_1) that the average site area is not independent of ecological zone. For this test, with data obtained by the sampling of sites, the population of sites in the Chevelon drainage would be enumerated and stratified according to ecological zone and a simple random

11. The need to take into consideration the power of a statistical test when working with small sample sizes has been discussed in G. Cowgill, "The Trouble With Significance Tests and What We Can Do About It," *AmAnt* 42 (1977) 350–368; and the concept applied to an archaeological example in V. Stanislowski, "A Rejoinder to Ackerly and Young's Comments on My Analysis of Formative Period Sites in the Valley of Mexico," *AmAnth* 87 (1985) 897–904.

12. A quadrat was selected by randomly picking the N–S and E–W coordinates of a point within the stratum and then selecting an orientation at random. The point was one corner of the quadrat, and the quadrat extended from that point in the direction of orientation.

sample of sites would be drawn from each of these three zones. The average site area would then be computed for each of the three samples and the null hypothesis tested using, say, a 1-way Analysis of Variance Test for difference in population means[13] (since more than two populations are being compared). No complications are introduced into the statistical analysis by virtue of the sampling procedure employed in this situation.

By formation of a site sample what is meant is a sampling procedure that does not use the site as the sampling unit, but instead identifies sites through sampling of another population made up of units with which sites are associated, such as a population of quadrats. Here, information on sites is recovered indirectly through examination of the sample units; for example, via the survey of quadrats selected as part of the sampling scheme for the population of quadrats. For the CARP survey the sample units were quadrats, and measures were made on these units—that is, the number of sites found in a quadrat, the predominant vegetation within the quadrat, the distance of the center of the quadrat to the nearest source of water, and so on. When the sample of sites is formed indirectly in this manner, the standard statistical inferential methods that assume the direct sampling of sites need to be modified. For instance, for the common problem of estimation of the number of sites in the region belonging to each class in a site classification, modification must be made of the variance estimates for the class proportions from what otherwise would be the variance estimate based on the binomial distribution if sites were to be sampled directly.[14]

Of the two sampling procedures, the first is preferable as it is directly amenable to the statistical analysis of site characteristics and, as noted by Cochran, "For a given size of sample [i.e., the total number of sites in the sample], a small unit [i.e., a quadrat which only contains one site] usually gives more precise results than a large unit [i.e., a quadrat which contains several sites]."[15] In other words, direct sampling of sites is preferable from an analytical point of view to sampling via quadrats which may contain several sites.

Since the enumeration of sites in a region is not possible, one can approximate a direct random sample of the population of sites by using quadrats sufficiently small in area so that the quadrat contains only one site or zero sites. The drawback is pragmatic: the much greater time cost for locating, surveying, and traveling among many small quadrats versus a smaller number of larger quadrats. The differences in time and cost affect the site sample size that can be obtained for a fixed time budget.[16] Consequently, the gain in precision that would otherwise be obtained by using small quadrats and locating the same number of sites as would be located with larger quadrats is lost when time constraints are considered. In addition, bias due to boundary effects are more pronounced with small quadrats, with their greater perimeter to area ratio, than is true for larger quadrats.[17] For the CARP survey, a rectangular quadrat of 100 m × 1000 m was chosen as giving a reasonable balance among these conflicting properties.

Quadrat Survey as Cluster Sampling

Recovery of sites via survey of quadrats can be treated as cluster sampling with unequal cluster sizes, as has been discussed elsewhere by several authors.[18] Associated with cluster sampling are a number of more specialized estimation procedures for population parameters and their standard errors. Essentially, estimation is based on treating the sampling unit (the quadrat) as the measurement unit by summing the values of the measure made on each entity (the individual site) contained in the sampling unit, and then using this summed value and the number of sample units (the number of quadrats) as the sample size for estimating the parameter in question and its standard error.

For example, one might be interested in estimating the average site area for all sites in the grassland zone. The procedure would be as follows. First, the area of each site found in the ith quadrat of the sample of quadrats would be measured and these site areas summed together to give a total site area, a_i, for the ith quadrat. These values would be used to compute an estimate, $\hat{A}$ for the parameter A (where A is the total site area summed over all sites in the stratum): $\hat{A} = \Sigma a_i$. Second, the estimated average site area would be computed from $\hat{A}/\hat{n}$, where $\hat{n}$ is the estimated number of sites in the region. Finally, the variance for the estimate $\hat{A}$ is computed using $N - 1$ degrees of freedom, where N is the total number of quadrats in the grassland stratum (as opposed to using d = [total number of sites in the sample] $-$ 1 as the value for the degrees of freedom).

The advantage of using cluster sampling for estimation purposes (as opposed to treating the sites as a sample and ignoring the cluster structure) arises when the

13. H. Blalock, *Social Statistics* (McGraw-Hill: New York 1972).

14. Cochran, op. cit. (in note 6) 64–68.

15. Ibid. 233.

16. Nance, 1981 op. cit. (in note 5).

17. Plog, op. cit. (in note 5).

18. Mueller, 1975 op. cit. (in note 5); S. Plog, 1978 op. cit. (in note 5); Nance, 1981 op. cit. (in note 5).

summed values for the measure (e.g., the values a_i) are 1) relatively homogeneous between clusters (since the variance estimate based on clusters is a between cluster variance estimate) and 2) the clusters are internally heterogeneous (since the variance estimate based on the value of the measure over each site is proportional to the sum of the variances within and between clusters).[19] When both 1) and 2) are true, the between cluster variance (that is, the variance of the summed values a_i) will be small in comparison to the variance in values as measured over each site. Hence, the standard error of the parameter estimates will be smaller when using the structure given by cluster sampling than would be the case if the site sample were treated as though it were obtained as a simple random sample of sites.

When translated into the archaeological context, this means that estimation procedures using the features of cluster sampling will be more precise (i.e., have a smaller standard error for the estimated parameter) when sites near one another (hence likely to fall into the same quadrat) are heterogeneous for the measure in question. The latter is valid for some measures (especially if a site and its nearest neighbor are not likely to be from the same phase and/or cultural context), but not reasonable for others. Situations where there is an advantage to using the more complex estimation procedures based on cluster sampling (with unequal cluster size) rather than treating the sample of sites as a single sample, undifferentiated by the method of site recovery, are thus context-specific. For this reason we will leave aside the methods of cluster sample estimation[20] and consider cluster sampling as a necessary epiphenomenon of the method of site discovery via quadrat survey, where the goal of the survey is discovery of sites in a representative and effective manner. It will be assumed that the data set created by the quadrat survey will be used for a variety of purposes. (Note that when the sample of sites obtained via quadrat survey is treated as if it were a simple random sample of sites, the variance estimate for a measure computed over the sites will be biased, but the bias is negligible as long as the number of sites is >50.)[21]

I use the term "site discovery" deliberately since regardless of whether or not the regional survey is to be done at a preliminary, discovery level (as was the case with the initial survey of the Chevelon drainage) with the data analyzed in an exploratory fashion (e.g., using

the methods of exploratory data analysis),[22] or in terms of specific hypotheses to be tested, sites must first be discovered before analysis may proceed. Discovery may be either general (what is the variety of sites in the region and in what contexts are they located?) or specific (what is the spatial distribution pattern of sites around point water sources such as a spring versus areal water sources such as a pond or lake?). Obviously, as data bases are developed, secondary or tertiary surveys may be oriented toward sites whose existence is already known, rather than toward the discovery of unknown sites. But the bulk of survey work is likely to be oriented toward the discovery of sites, whether for general or specific purposes.

Once we view the survey as oriented toward creating a sample of sites, we may focus further considerations on the properties that any sample should satisfy, namely, representativeness of the population from which the sample is drawn and adequacy of the sample for efficient estimation of population parameters.

Representative Sampling of Sites

A representative sample can be defined through the relationship between a random sample of population elements from a population and the characteristics of that population. In direct simple random sampling where the sampling unit is a population element, each of the elements is equally likely to be selected. Hence the relative frequencies associated with any classification of the elements comprising the population will be preserved in the sample in the sense that the expected relative frequencies for the classes in the classification are precisely the population relative frequencies (i.e., the sampling procedure is unbiased). To put it another way, estimates of relative frequencies measured over random samples from the population will average out "in the long run" to exactly the population parameters for which they are the estimates. Thus, a sample will be said to be representative of a population if the sampling procedure used to obtain the sample is unbiased in the sense just described.

Under this definition, then, the procedure for creating a sample of sites through survey of a simple random sample of quadrats as the sampling units also forms a representative sample of sites. Regardless of the spatial

19. Cochran, op. cit. (in note 6) 208.

20. See Nance, 1981 op. cit. (in note 5) for an extensive discussion as it applies to archaeological data, and Cochran, op. cit. (in note 6) for the mathematical properties.

21. Cochran, op. cit. (in note 6) 239.

22. J. Tukey, *Exploratory Data Analysis* (Addison-Wesley: Reading, MA 1977); see also Clark, op. cit. (in note 2); C. Carr, "Getting Into Data: Philosophy and Tactics for the Analysis of Complex Data Structures," in C. Carr, ed., *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Technique, and Theory* (Westport Press: Kansas City 1985) 18–44; D. Read, "Statistical Method and Archaeological Theory: Enlightenment Out of Discordance?" in Carr, ed., op. cit. (in this note) 44–91.

distribution of sites in the region, the randomness of the location of the survey units with respect to the spatial patterning of sites carries over to the sites within the quadrats. Thus the expected relative frequencies for site class frequencies will be the relative frequencies for those classes in the population of sites in the region. Further, as long as the sampling procedure (e.g., stratified random sampling, systematic unaligned sampling, etc.) does not introduce correlations between features of site spatial distribution and quadrat selection (as may occur with simple systematic sampling), a representative sample of sites will still be produced. Contrariwise, so-called purposeful sampling will not generate a representative sample of sites as can be theoretically demonstrated and as has been shown empirically,[23] unless the site classes are randomly distributed in space, a condition which belies any rationale for purposeful sampling.

Given that simple random sampling of quadrats is easily defined and will produce a representative sample of sites, one might well ask why more complicated sampling programs should be introduced. The answer lies in sampling effectiveness (here to be taken as the number of sites that can be recovered for a given amount of resource expenditure) and its impact on the precision of a parameter estimate.

Effective Sampling of Sites

For the purposes of this section it will be assumed that a decision has already been made as to the size (see below) and form of a quadrat (see above), so that resource expenditure is directly related to the number of quadrats to be surveyed. It will be further assumed that alternative sampling schemes have essentially the same average cost per quadrat, so that the number of quadrats surveyed can be taken as a proxy measure for resource cost regardless of the sampling scheme. Here we will only be considering quadrat sampling, hence all sampling procedures will refer to the method of selection of quadrats, not of sites. With these constraints, different quadrat sampling schemes may be compared by keeping the number of quadrats fixed and determining the expected number of sites to be recovered under each scheme.

The number of sites that can be recovered for a given expenditure of resources is used here as a measure of sampling effectiveness due to the close relationship between sample size and the precision of an estimator.

What is meant by an "estimator" is the procedure (such as computing the sample mean for a random sample of a given size) used to obtain an estimate of a population parameter. The precision of an (unbiased) estimator is measured by the variance of estimates measured across different samples obtained in the same fashion from that population, keeping the sampling procedure (including the sample size) fixed. The precision of an (efficient) estimator is, in general, inversely proportional to the sample size.[24] Thus, under simple random sampling the precision of an estimate based on an estimator using larger sample sizes is greater than one based on smaller sample sizes.

Precision is also affected by the sampling procedure. For example, cluster sampling where the sampling unit (such as the quadrat) contains several of the population elements (in this case, sites) will be more precise than simple random sampling for the same total number of population elements in a situation where the between-cluster variance is low and the within-cluster variance is high, as previously discussed.

For the reasons outlined in the previous section, it will be assumed that the sample of sites produced by the survey of quadrats will be treated as though it were obtained as an element sample of sites, even though in specific situations this will not always lead to the most precise estimate possible. Under these conditions precision is directly related to the number of sites recovered, and this justifies the use of the total number of sites recovered as a measure of sampling effectiveness.

Recovery of Site Types through Simple Random Sampling of Quadrats

An initial topic addressed by the CARP project was discovery of the variety of sites and their frequency in the Chevelon drainage. Site type frequencies are estimated from the sample data, but the variety of site types that are discovered depends on the sampling effectiveness. In general there will be a direct relationship between the number of sites recovered and the number of site types in the sample, with the latter "asymptotically" reaching its true value as the sample size increases. When there is *a priori* knowledge of the frequency distribution of site types, it is possible to estimate the total number of site types in the population,[25] and confidence intervals for the estimated number of site types may be

23. D. Brose, "Locational Analysis in the Prehistory of Northeast Ohio," in C. Cleland, ed., *Cultural Change and Continuity* (Academic Press: New York 1976) 3–18; D. Alexander, "The Limitations of Traditional Surveying Techniques in a Forested Environment," *JFA* 10 (1983) 177–186.

24. R. Elandt-Johnson, *Probability Models and Statistical Methods in Genetics* (John Wiley and Sons: New York 1971).

25. E. Pielou, *An Introduction to Mathematical Ecology* (John Wiley and Sons: New York 1969) 208–214.

constructed[26] to give a measure of the extent to which it is likely that all site types have been recovered. For the CARP project, however, such *a priori* data were not available.

In the absence of these data, we may examine the likelihood that the sampling program has recovered instances of all the site types by considering the likelihood that at least one site of a site type has been found. This may be done by computing the quadrat sample size necessary to ensure, with a specified degree of certainty, that at least one of a specific site type in the population of all sites is included in the sample of sites obtained from the survey of the quadrats. For the purpose of this argument it will be assumed that a simple random sample of quadrats has been selected.

Consider the population of sites to have been divided into classes as follows. Suppose there are N sites in the region and these have been classified (or would be so classified if one had the complete collection of sites) into m classes which are labeled $c_1, c_2, \ldots, c_m$. Denote the number of sites in the kth class, c_k, by n_k, so that the relationship $n_1 + n_2 + \ldots + n_m = N$ holds. Suppose the region is divided into Q quadrats. Agree that one wants to be, say, $1 - \alpha$ confident that in the sample of sites obtained by randomly selecting q of the Q quadrats there will be at least one site from each of the m classes somewhere in the q quadrats. This symbolism provides a characterization of the question posed above.

Simplify matters by examining only the class with the fewest number of sites. Call this class c_{min}, and suppose it has n sites. How many quadrats must be selected to be $1 - \alpha$ confident that at least one site from c_{min} is in the sample of quadrats? Simplify still further by assuming (since only the rarest sites are being considered) that a quadrat will have either zero or one site from the class c_{min}. That is, assume that the sites of the same type are spatially dispersed on a scale greater than that of the dimensions for the quadrats. Thus, there will be n quadrats which have one site each from c_{min} and $q - n$ quadrats without such a site. The goal is to establish the probability that the random selection of q quadrats will have included at least one quadrat with a site from c_{min}.

To have at least one site means that one has failed to select only quadrats which do not have such a site. If one chooses a quadrat randomly, the likelihood, p, of it *not* containing a site from c_{min} is given by $p = ([Q - n]/Q)$, so that the likelihood $L = 1 - \alpha$ of having at least one site from c_{min} somewhere in the q quadrats is given by $1 - p^q = 1 - ([Q - n]/Q)^q$.[27]

Set $1 - \alpha = L = 1 - ([Q - n]/Q)^q$
and solve for q to obtain:
$$q = \log (1 - L)/(\log [Q - n] - \log Q). \tag{1}$$

Now use Equation (1) to examine three questions: 1) how many quadrats must be surveyed to ensure finding a site of a given rarity?; 2) what is the rarest type of site, from a specified list of site types, that one is likely to find in a given sampling fraction of the population of quadrats?; and 3) given the number of quadrats that will be surveyed and the rarest type of site to be found, what percentage of the total area of the region should each quadrat represent?

First, consider some values for the terms in Equation (1). Suppose $\alpha = 0.05$, $Q = 100,000$ (e.g., the region contains 10,000 sq km and the quadrats are of area 0.1 sq km. The choice of 10,000 sq km and quadrats of area 0.1 sq km is based on the CARP survey.) Set $n = 10$. If these values are substituted into Equation (1), it follows that $q = 32,000$. That is, one needs to survey approximately one-third of the region to have a 95% chance that a site type only represented 10 times will be included in the sample of sites created through a random sample of quadrats.

For the second question, suppose that there is to be a 1% sample of the region; i.e., the sampling fraction q/Q is 0.01. For a given type of site, with what frequency must it occur in order that it will be found in at least one quadrat with, say 95% certainty? Substitute $L = 0.01$, $Q = 100,000$ and $q = 0.01 \times 100,000 = 1,000$ into Equation (1) and solve for n to obtain $n = 200$.

Finally, consider the situation where one is willing to survey 100 quadrats and wants to be 95% confident that a site whose type is represented only 10 times is included. How large should the quadrats be? If these numbers are substituted into Equation (1) and Q is solved for, the result is that $Q = 400$. Hence, each quadrat should contain 100,000/400 or 250 sq km.

These results suggest three significant limitations of simple random sampling of quadrats over the region. First, ensuring the inclusion of even moderately rare sites can be achieved through simple random sampling of quadrats only at the expense of having to take a prohibitively large sampling fraction.[28] On the basis of the CARP survey, $q = 32,500$ quadrats translates into at least 15,000 crew days of surveying—a prohibitively large time investment. Second, a more plausible sampling fraction is highly unlikely to find a rare site type

26. K. Kintigh, "Measuring Archaeological Diversity by Comparison with Simulated Assemblages," *AmAnt* 49 (1984) 44–54.

27. Strictly speaking, this equation is not correct since it assumes

sampling with replacement and the sampling of quadrats is done without replacement. Pragmatically, however, the error introduced is small unless the value for g is close to Q, which is unlikely.

28. Nance, 1981 op. cit. (in note 5).

specified in advance.[29] And third, adequate coverage of the region by a reasonable number of quadrats can be achieved only by increasing the size of the quadrats to a point where the notion of a quadrat being a sample of a small segment of relatively homogeneous space has become meaningless.

Regional Stratification

Now let us consider means for increasing the effectiveness of the original sampling scheme of simple random selection of quadrats. Two situations will be examined: first, no *a priori* information is available on the spatial distribution of sites and, second, *a priori* information is available on the distribution of sites.

As a base consider the situation where one is unwilling, *a priori,* to make any predictions about site location and site density for a region. In other words, the possibility of a random spatial distribution of sites across the region is not excluded. For a spatially random distribution of sites, any alteration of the simple random sample is unjustified in terms of the criterion of sampling effectiveness. Hence, for a region or subregion in which one is unwilling to make any *a priori* statements about site location, the best sampling scheme is a simple random sample of quadrats.

In most instances, however, there are reasons for arguing that a random spatial distribution does *not* characterize the site distribution in the region.[30] For example, when the initial sampling design was formulated for the CARP survey, though little was known about site distribution in the Chevelon drainage, limited surveys by other researchers and information on site distribution in nearby areas such as Hay Hollow Valley suggested that a stratification by ecological zone would also be a stratification of sites by site density.[31]

Site Density Estimation in Low Density Subregions

The Chevelon drainage has a clear division into grassland, juniper-piñon, and ponderosa pine zones. These also represent increasing elevation from around 4000 ft at the Little Colorado River to around 6000–8000 ft at the Mogollon Rim. Initial expectations on site density were qualitative. The ponderosa pine zone was thought to have few, if any, sites; sites should be plentiful in the juniper-piñon zone; and the density of sites in the grassland should be somewhere in between the site density for the two other zones. The ponderosa pine zone represents about one-half of the total area and the remaining area is split roughly with three-fifths in the juniper-piñon zone and two-fifths in the grassland zone.

Assume, to make the numbers simple, that in fact there are 100,000 sites in the drainage (the CARP surveys indicated that this figure is probably correct to one order of magnitude). Further, assume that time and resources limit surveying to 100 quadrats (which is approximately the number surveyed during the initial fieldwork). These data give an overall average density of 10 sites/sq km, or 1 site/quadrat (of area 0.01 sq km), and thus an expected total of 100 sites to be found by simple random sampling of quadrats over the whole drainage. Elimination of the ponderosa pine zone decreases the region to 5000 sq km, yielding an average density of two sites/quadrat in the remaining area and hence a doubling of the expected number of sites to be found to 200. The simple (but perhaps drastic) expedient of eliminating a subregion believed to be void of sites has doubled the effectiveness of the sampling.

Before considering the differential density of sites in the two remaining zones, the possibility that there may in fact be sites in the ponderosa zone needs to be examined further. Here one is faced with the difficult choice between drastically decreasing the total number of sites to be discovered with the 100 quadrats of the survey in order to find rare sites (if they exist) in one subregion, versus eliminating a zone from the sampling universe which may in fact contain sites. If there are, say, a total of 10 sites in the ponderosa pine zone, then, as previously argued, one would need to survey about one-third of that region, or approximately 10,000 quadrats, to have 95% confidence that the sampling program would find even one of these 10 sites.[32] Thus, to ensure finding even one rare site requires a number of quadrats two orders of magnitude larger than that allocated for the survey program. This suggests that one has little choice but to delete the ponderosa pine zone from the probabilistic sampling program if there are but a few sites in that zone.

On the other hand, one would certainly want assurance that the site density is low before excluding the ponder-

---

29. Note that this statement only applies to a single site type distinguished in advance. Obviously, if each site in a region is unique, then many rare site types will be found. For the situation in which site types are not known in advance, a single class, Rare Sites, can be defined and the above results would apply to that class. Thus for the example given above, the class Rare Sites would have to contain at least 300 members in order that there be 95% certainty that at least one such site would be found. Most of the remaining 299, which are likely to represent a diverse group of site types, would not be found, thus leaving many site types undiscovered.

30. Binford, op. cit. (in note 5).

31. F. Plog, personal communication 1971.

32. Possibly one may not need to survey all 10,000 quadrats since such a site might be found prior to examination of the last quadrat.

osa pine zone. The argument for low site density in this subregion is inductive and the extension of data from other areas may not be valid. The assumption of low site-density may be tested by surveying a small number of quadrats and, if none of these quadrats includes sites, estimating the maximum density of sites consistent with those quadrats not having any sites. Note that the usual maximum likelihood estimate for site density is not useful here since the estimate would be zero if the quadrats contain no sites. However, the probability of a Type II error is high if this estimate is accepted when few quadrats are surveyed, so an alternative procedure aimed at reducing the probability of a Type II error is given here.

The author suggests the following procedure for estimating the maximum site density consistent with finding no sites in a survey of k quadrats. The method consists of a test for the following null and alternative hypotheses:

H_0: Site density $D \leq D_0$;
H_1: Site density $D > D_0$;   (2)

where D_0 is a hypothesized value for site density in the ponderosa pine zone. The null hypothesis, H_0, is accepted if none of the k quadrats contains sites, and rejected if at least one quadrat does contain a site. A value is selected for D_0 so that Prob[Type I error] = α, where α is the significance level. The estimate of the density D, namely $\hat{D}$, is set equal to the value of D_0.

The value for D_0 is obtained by noting that it must satisfy the condition that Prob [0 sites in k quadrats] = $1 - \alpha$. The latter probability may be computed as follows. The probability of a randomly selected quadrat not containing a site is $1 - a D_0$, where a is the area of the quadrat.[33] Hence Prob [0 sites in k quadrats] = $(1 - a D_0)^k$. Set the latter expression equal to $1 - \alpha$ and solve for D_0 to obtain

$$D_0 = (1 - [1 - \alpha]^{1/k})/a = \hat{D}. \quad\quad (3)$$

For example, if $\alpha = 0.05$, $k = 2$, no sites are found in either quadrat, and $a = 0.1$ sq km (as was the case for the CARP survey), then $\hat{D} = (1 - 0.95^{1/2})/0.1 = 0.25$, so an estimated density of $\hat{D} = 0.25$ sites/sq km is taken as the estimate of the upper bound on the value of D. This corresponds to about 1250 sites in the 5000 sq km of the ponderosa pine zone, in comparison to the estimated 100,000 sites in the juniper-piñon and grassland zones, or a two orders of magnitude difference in estimated number of sites. Whether one should survey

further in the ponderosa pine zone depends on the weighing by the archaeologist of the need for additional information relating that zone to the pattern of site distribution in the drainage as a whole against the competing need for effective sampling of the other two zones. For the CARP survey, the decision favored more effective survey of the juniper-piñon and grassland zones when no sites were found in the two quadrats surveyed in the ponderosa pine zone.

The example shows that at little cost to the allotted budget of 100 quadrats, the validity of the assumption that the site density in the ponderosa pine zone is low can be (and was) tested. The estimates $\hat{D}$ for the density D corresponding to a range of values for k, the number of quadrats surveyed, and a, the area of a quadrat, are given in Table 1.

Sampling of Strata with Unequal Site Density

Consider next the unequal density of sites in the grassland and the juniper-piñon zones. A difference in density on the order of a factor of 10 was considered reasonable (or a density estimate could be made in a preliminary survey). This translates into a density for the grassland of about 3.1 sites/sq km and a density for the juniper-piñon zone of about 31 sites/sq km, given 100,000 sites for the two zones.

If the quadrats are allocated through proportional allocation based on the area of the zones, then about two-fifths, or 40, of the quadrats would be in the grassland and about three-fifths, or 60, of the quadrats would be in the juniper-piñon zone. Were the quadrats so placed, one would expect to find $40 \times 3.1 \times 0.1 = 12$ sites in the grassland and $60 \times 31 \times 0.1 = 186$ sites in the juniper-piñon zone, for a total of 198 sites.[34] Thus the expected number of sites is precisely what would be observed if the quadrats had been placed randomly over the two subregions without stratification. Hence allocation proportional to strata area gives no increase in sampling effectiveness. But with unequal density of sites in the two strata, the standard error for estimates of the total number of sites in each stratum will differ, indicating that one stratum is relatively oversampled and the other is relatively undersampled for the purposes of estimating the total number of sites.[35]

---

33. It is being assumed that site density is sufficiently low and/or sites are sufficiently spatially dispersed so that there will be at most one site per quadrat. With higher site density and/or more clustered distribution of sites, the quantity $1 - aD$ will be an underestimate of the probability that a quadrat contains no sites (see Nance, 1983 op. cit. [in note 5] 312–316.

34. The deviation from the 200 sites expected under simple random sampling is due to round-off error.

35. Over- or undersampling is measured by the variance of a parameter estimate in the different strata and depends on the variability of the measure in question for each stratum. Measures other than the number of sites/quadrat may have a different variance pattern for the strata. The focus here and below has been on the number of sites/quadrat as this is a basic measure common to most surveys.

Number of Quadrats	Area of a Quadrat (sq km)				
	0.05	0.1	0.2	0.5	1.0
Significance Level $\alpha = 0.05$					
1	1.00	0.50	0.25	0.10	0.05
2	0.50	0.25	0.125	0.05	0.025
3	0.33	0.17	0.085	0.03	0.017
4	0.25	0.125	0.063	0.025	0.0125
5	0.20	0.10	0.05	0.02	0.01
10	0.10	0.05	0.025	0.01	0.005
15	0.07	0.03	0.017	0.006	0.003
Significance Level* $\alpha = 0.10$					
1	2.00	1.00	0.50	0.20	0.10
2	1.00	0.50	0.25	0.10	0.05
3	0.67	0.33	0.17	0.06	0.03
4	0.50	0.25	0.125	0.05	0.025
5	0.40	0.20	0.10	0.04	0.02
10	0.20	0.10	0.05	0.02	0.01
15	0.13	0.07	0.03	0.012	0.006

*10% significance level is more conservative with respect to density estimates consistent with no sites found in the surveyed quadrats.

Table 1. Site density estimate $\hat{P}$ for a sample of K quadrats each containing zero sites. Density equals number of sites per sq km.

Correction may be made by allocating quadrats in accordance with Neyman optimization[36] which, in this context, implies (assuming the resource cost per quadrat is the same in all strata):

$$n_{JP}/n_G = (A_{JP}/A_G) \times (D_{JP}/D_G), \qquad (4)$$

36. The differences between Neyman optimization and sampling for an equal number of sites in each stratum should be carefully noted. Neyman optimization will be more efficient for estimating region-wide parameters such as the total number of sites in the region, but at the cost of decreasing the number of site types that are likely to be found—for example, in the grassland stratum, since the expected number of sites to be recovered from that stratum is 5.6 sites. In contrast, sampling for an equal number of sites in each stratum should recover a greater proportion of the site types in the grassland stratum since the expected number of sites to be found is now five times as great as under Neyman optimization. But this is achieved only at the cost of reducing the total number of sites in the region to be found, hence making it less likely that as many of the site types in the region will be recovered, and making statistical tests of hypotheses less powerful. Which of these two sampling programs (or even a third strategy) is best cannot be answered via sampling theory as this requires an archaeological decision to be made about the type of information that needs to be recovered. The latter, ultimately, provides the basis for deciding upon a sampling strategy. Sampling theory provides effective and efficient sampling techniques for specified goals in the context of any constraints (such as having to use cluster sampling due to the nature of data recovery) that must be satisfied. The goals require archaeological definition.

Unfortunately the differences among sampling strategies are not minor as this example illustrates; e.g., an expected recovery of 260 versus 56 sites for the same number of quadrats that will be surveyed. Thus it is crucial that the archaeological goals be well identified before committing a project to a particular sampling design.

where:

n_{JP} = number of quadrats placed in the juniper-piñon stratum,

n_G = number of quadrats placed in the grassland stratum,

A_{JP} = area of the juniper-piñon stratum,

A_G = area of the grassland stratum,

D_{JP} = site density in the juniper-piñon stratum, and

D_G = site density in the grassland stratum.

For the CARP survey, n_{JP}/n_G = (3000 sq km / 2000 sq km) × (31 sites/sq km)/ (3.1 sites/sq km) = 4.75. For a sample size of 100 quadrats it follows that n_{JP} = 82 and n_G = 18. The expected numbers of sites are now 18 × 3.1 × 0.1 = 5.6 sites for the grassland and 82 × 31 × 0.1 = 254 sites in the juniper-piñon for a total of about 260 sites versus the 198 expected under (areal) proportional sampling. Allocation by optimal placement has increased the expected number of sites to be found by about 30% while simultaneously ensuring that no stratum is over- or undersampled as measured by the variance of the estimate of the number of sites in each stratum.

These three sampling strategies—simple random sampling of all quadrats in the region, stratification into zones with few or no sites and zones with sites, and optimal allocation of quadrats—have yielded, in this example, expected values of 100, 198, and 260 sites, respectively. Utilization of a minimal amount of information about the region has increased the expected effectiveness of the survey by a factor of ca. 2.6 when

simple random sampling is compared with optimal allocation.

These results presume that there exist no separate, overriding archaeological issues regarding allocation of quadrats. Any such archaeological issues would obviously take precedence. For example, if the goal were to have an equal number of sites from the grassland and from the juniper-piñon zones, then it would be necessary to allocate about 91 quadrats to the grassland and about 9 quadrats to the juniper-piñon zone, for an expected number of $91 \times 3.1 \times 0.1 = 28$ sites and $9 \times 31 \times 0.1 = 28$ sites respectively, or a total of 56 sites for the two zones. This sample size will be about 20% as efficient for statistical testing as is the sample of 260 based on Neyman optimization.[37]

Sampling Precision

The above results are stated in terms of expected numbers of sites to be discovered and represent what will be found on the average if one were to repeatedly sample regions with the same characteristics and the same sampling program. In any particular survey, the actual number of sites found is likely to deviate from the expected value, with the magnitude of the deviation having a probability distribution dependent on the precision of the estimator; that is, on the variance of the estimates around the expected value of the estimates (which will equal the population parameter in question for an unbiased estimator). The precision of the estimator for a given sample size will be affected by the characteristics of the spatial distribution of sites. The actual number of sites recovered will tend to be closest to the expected value (and hence there will be greater precision) for a uniform distribution and subject to greatest deviation when the sites are spatially clustered. That is, greater deviation (or reduced precision) is likely when the number of sites per quadrat has larger variance. The latter will also be affected by the size of the quadrats. Consequently, the effect of the size of quadrats on the precision of an estimator also needs to be examined.

Though some authors have measured the precision of an estimator experimentally[38] through use of a natural data set with known location of sites, this method has limited usefulness since the results cannot be generalized to other regions with different spatial distributions of sites. Instead, we will take advantage of the theoretical work that has been done on this topic.

Sampling procedures (in this case, sampling with quadrats with small areas versus quadrats with large areas) may be compared to one another by computing their relative efficiency, or ratio of the respective variances associated with each sampling procedure, when each is used to estimate the same population parameter. For example, if one sampling strategy is simple random sampling of the elements of a population and the second sampling strategy is cluster sampling of the same population using spatially defined sampling units, then it can be shown that the relative efficiency of these two sampling strategies varies from favoring simple random sampling of elements to favoring cluster sampling as the distribution of the elements in the population varies from a uniform to a clustered distribution, holding constant the number of elements, M, in each cluster in the cluster sample.[39] This implies that the relative efficiency of small quadrats in comparison to large quadrats (holding other factors constant) varies inversely with the spatial distribution of sites as that distribution varies from uniform to clustered. In other words, small quadrats (defined as quadrats for which there will be either one or zero sites per quadrat) are generally more efficient in the sense of yielding more precise estimates of the number of sites in the region than are large quadrats (defined as quadrats for which there will either be zero or several sites) when sites are spatially clustered. Read[40] shows by example that the difference in precision can vary by a factor of three between small and large quadrats when comparing uniform to clustered spatial distributions.

With large regions and subregions, sites are likely to be spatially clustered; hence smaller quadrats will tend to yield more precise parameter estimates than with larger quadrats. For a quadrat of a given areal size, the expected number of sites per quadrat is proportional to the average density of sites. Consquently the precision of an estimate of, for instance, the total number of sites in the region, based on a fixed total area to be surveyed via quadrats, can be controlled by changing quadrat size in accordance with the average site density within a stratum.

This procedure suggests that sampling will be most effective when an initial sample is used to define the characteristics of the spatial distribution of sites in strata based, for instance, on ecological and topographic properties, and then these strata are stratified further by the location(s) of site clusters. To put it another way, the spatial clustering of sites within a stratum indicates that the stratum is characterized by subareas with high and low densities of sites. As argued above, sampling of

37. Cochran, op. cit. (in note 6) 98–99.

38. S. Plog, "Relative Efficiencies of Sampling Techniques for Archaeological Surveys," in K. Flannery, ed., *The Early Mesoamerican Village* (Academic Press: New York 1976) 136–158.

39. Cochran, op. cit. (in note 6).

40. Read, 1975 op. cit. (in note 5).

such a stratum is made more effective for a given total area to be surveyed (in terms of recovering a greater number of sites) by restratifying according to areas of high density and low density, and allocating the number of quadrats per substratum using optimal allocation which depends on the relative site density per substratum. Hence restratification by areas of high and low density and optimal allocation of quadrats will simultaneously increase the effectiveness of the sampling and help ensure that the actual number of sites recovered is close to the expected number for that sampling program. Quadrat size (and hence the number of quadrats, keeping fixed the total area to be surveyed) may be chosen (within the constraints imposed by the time required to locate and travel to quadrats) so that within each stratum the expected number of sites per quadrat is close to one if site density within a stratum is not approximately uniform.

To see the effect obtainable from restratification by site density, consider one part of the Bureau of Land Management survey of the California desert mentioned in this article's introduction. A survey of quadrats in a region varying topographically from mountainous to desert valley (divided into two strata: 1) Mountain and 2) Valley Floor) indicated that within the Valley Floor stratum, sites tended to be clustered either along the break in the gradient at the edge of the valley, or in the lowest part of the valley, with few sites in between.[41] Hence the stratification into Mountain versus Valley Floor can be made more effective, and estimates of site number more precise, by stratifying the valley area into, say, "Valley Bottom," "Valley Edge," and "Other," and optimally allocating sample units using the estimated site density in these strata obtained from the initial survey. Since the "Other" stratum comprises perhaps three-fourths of the valley floor and has a very low site density, optimal allocation of quadrats should increase the expected number of sites to be found by at least a factor of three.

Sampling in Low Density Strata

If, after restratification, one is unwilling to exclude a stratum with low site density, then one needs to know the number of quadrats that must be assigned to such a stratum to ensure that at least one site is found. That number may be determined as follows. Suppose the ponderosa pine zone (to return to the CARP situation) is thought to contain enough sites to warrant its inclusion in the probabilistic sampling program, and say that the site density is thought to be at least one site/sq km in

that stratum. We want to find the number, m, of quadrats that must be surveyed so that there is a $1 - \alpha$ probability of finding at least one site.

The number of quadrats, m, is computed as follows. For each quadrat of area, say, 0.1 sq km, the probability, p, that it contains no sites is (approximately) given by $p = 1 - $ (site density) $\times$ (quadrat area) $= 1 - 1.0 \times 0.1 = 0.9$. Thus to have, say, a $1 - \alpha = 0.95$ probability of finding at least one site after sampling m quadrats, one needs to have $0.95 = 1.0 - (0.9)^m$. Solving for m yields $m = 28$. Thus, if 28 randomly chosen quadrats are surveyed and the density is $\geq$ one site/sq km, at least one site will be found with 95% confidence. If no sites are found it may be concluded (with 95% confidence) that the density is less than one site/sq km.

Note the difference between this situation where one wants to ensure that at least one site will be found, given a lower bound on the site density, D, and the previous situation where an estimate was made of an upper bound for that site density. In the CARP survey, the two quadrats without sites implied that an upper bound on the site density would be 0.25 sites/sq km, or well below the value of 1.0 sites/sq km used in the above example. In other words, even with an unrealistically high estimate of 1.0 sites/sq km in the ponderosa pine zone (which is inconsistent with the fact that no sites were found in the two quadrats that were surveyed,[42] over one-quarter of the allotted 100 quadrats would be needed to find even one site in that zone. And if the estimated upper bound of 0.25 sites/sq km for the site density in the ponderosa pine zone based on two quadrats devoid of sites is used, a minimum of 118 quadrats, or more than the budgeted 100 quadrats, would need to be surveyed to have an 0.95 probability that at least one site would be found. While the option of excluding a region from the probabilistic sampling program may seem drastic, these data indicate that pragmatically there is little alternative. In this situation purposeful sampling may be more useful than probabilistic sampling if the goal is simply to locate sites in a low density area. Table 2 gives the number of quadrats that need to be surveyed to ensure finding at least one site for a range of density values.

Sampling for Spatial Patterning

It has been shown that effective sampling requires stratifying a region by site density and site clustering.

41. G. Coombs, *The Archaeology of the Northeast Mojave Desert* (Bureau of Land Management: Sacramento 1979) 88, 91.

42. If the density were one site/sq km, then the probability that two quadrats had no sites would be given by $1-(0.9)^2 \approx 0.20$. Thus the null hypothesis, $H_0 : D \geq 0$, may be rejected at the 20% significance level (assuming a willingness to accept a fairly high chance of making a Type I error in order to increase the power of the test) in favor of the alternative hypothesis, $H_1 : D < 1$, again confirming the "guess-timate" of a low site density in the ponderosa pine zone.

Table 2. Minimum number of quadrats for ensuring discovery of at least one site with probability $1 - \alpha$.

Site Density*	Area of a Quadrat (sq km)			
	0.05	0.10	0.20	0.50
Significance Level $\alpha = 0.05$				
1.0	58	28	13	4
2.0	28	13	6	1
3.0	18	8	3	1
5.0	10	4	1	1
Significance Level $\alpha = 0.10$				
1.0	45	22	10	3
2.0	22	10	5	1
3.0	14	6	3	1
5.0	8	3	1	1

*Density equals number of sites per sq km.

These properties relate to the spatial distribution of sites as the remnant of a collection of settlements forming the spatial nexus for a past social system. Depending upon the details of a model for such a system of interrelated settlements, a variety of spatial distribution patterns are expectable, ranging from those in which settlement location is largely responsive to local resources, (such as !Kung San camps being located near waterholes in the dry season)[43] to settlement patterns primarily affected by the system of group interaction (such as the location of retail markets as predicted by Central Place Theory).[44]

While it is not possible to use a single spatial pattern as characteristic of all settlement systems, common to these various patterns is the feature of having, over a sufficiently large scale, subareas of relatively high density of settlements and subareas of relatively low density of settlements. One can, as a first approximation, equate areas of high density with a single subsystem, or possibly a limited number of subsystems. But because the spatial distribution of sites may also represent a compression of time, the equivalent statement for sites is less certain. Nonetheless, differential density and clustering of sites represents a minimal feature of spatial distributions that can be related to the structure of settlement systems, as has been discussed by Hodder and Orton.[45]

This suggests the construction of a second sampling strategy that is effective for outlining the spatial distribution of sites in the region, as defined by differential site density. Two general issues arise in this type of sampling: first, location of areas of relatively high density of sites along with a measure of that density and, second, determination of the boundaries of such areas.

Effectiveness in locating areas of relatively high density is related to the sampling intensity of the population of quadrats, since areas of higher site density can be determined in an initial survey through the frequency distribution of the number of sites per quadrat. An area of higher density is defined by having several quadrats in close proximity, where each quadrat contains one or more sites. Hence the sampling intensity should be great enough to ensure that several quadrats will intersect an area of greater site density. The density of sites in these quadrats is an estimate of the site density for that area and the location of the quadrats is obviously the location of the area.

Boundaries of areas of higher density of sites can be found through systematic sampling (in contrast to simple random sampling within strata for parameter estimation). Given a quadrat that is believed to be located within a higher density area, place additional quadrats at regular intervals along perpendicular axes (e.g., N-S and E-W) intersecting at the given quadrat until the site density per quadrat decreases markedly. The distance between quadrats would depend upon the expected dimensions of the area of higher site density and the accuracy with which the boundary is to be located. This procedure was used effectively in the Chevelon sampling program.

Within areas of high site density so outlined, 100% surveys may be necessary for defining internal spatial patterning of site distributions, as defining the spatial pattern of site distribution at a fine level is difficult when employing a random sample of quadrats. A pattern is not a population and statistical inference from a sample to a population does not hold, in general, for spatial patterns. Unsurveyed areas require extrapolation from surveyed areas to fill out the pattern. In contexts where

43. J. Yellen, and R. Lee, "The *Dobe-/Du/da* Environment," in R. Lee and I. DeVore, eds., *Kalahari Hunter-Gatherers* (Harvard University Press: Cambridge 1976) 27–46.

44. Smith, ed., op. cit. (in note 9).

45. Hodder and Orton, op. cit. (in note 9).

the pattern can be approximated by a continuous distribution with relatively smooth changes in the distribution across space, statistical interpolation techniques such as trend surface analysis can be used.[46] But the patterning of the spatial distribution of sites is discrete and thus only approximated by a continuous distribution. Hence a 100% survey may be needed to define the spatial distribution of sites at a fine level. While a complete survey of the region may be impossible, limited areas can be so surveyed.[47] The utility of complete surveys of subregions depends, however, upon first establishing the overall pattern of site distribution in the region.

These two strategies—sampling of quadrats to construct a sample of sites to be used for analysis and sampling for determining spatial patterning—can be integrated into a multi-stage sampling program.[48] Effective sampling of sites requires data from site spatial distributions, so the first stage of the sampling program can be aimed at defining the general characteristics of the spatial distribution of sites. Subsequent stages can be aimed at sampling based on the information so gained, either for parameter estimation, or for definition of the spatial distribution at a fine level.

Conclusion

Regional sampling is a technique for obtaining a broad data base which will be used for a variety of purposes. The goal is not just the standard survey sampling goal of precise parameter estimation, but to maximize the total amount of information that can be recovered under the constraint that the collection of sites (found through the survey of quadrats) is to be representative of the whole collection of sites.[49] The primary means for increasing the effectiveness of the sampling design, as measured by the number of sites discovered for a given number of quadrats, has been shown to be fine-grained stratification of a region that includes relative density and spatial clustering of sites as part of the criteria for defining strata. Fine-grained stratification of a region may require strata consisting of noncontiguous segments of space.[50]

Cognizance should also be made of an issue that has been raised a number of times against probabilistic sampling of regions, with probabilistic sampling taken as simple random sampling. Namely, rare sites that can be discovered by other means (e.g., they are large and obvious) will be missed. The criticism has limited validity in the sense that simple random sampling of quadrats is likely to miss rare sites, regardless of their importance for descriptive and explanatory purposes. But this is not valid as an argument against sampling procedures as discussed here, since regional sampling is not identical to simple random sampling of quadrats. The rare, large site owes its archaeological importance in part to its role as a center for settlement interaction. Consequently, that large site is part of the data to be included in sampling for the spatial distribution of sites as reflecting the pattern of settlement interaction. As long as there is control over the relationship between the characteristics of the sample of sites and the population of sites, it is possible to make independently replicable, statistically, and archaeologically sound inferences about that whole collection of sites. A well-designed probabilistic sampling program is an effective means for defining and controlling that relationship while simultaneously providing a substantial data base for the research interests of the archaeologist.

---

uation of Sampling Strategies: Simulated Excavations of a Kenyan Pastoralist Site," in I. Hodder, ed., *Simulation Studies in Archaeology* (Cambridge University Press: Cambridge 1978) 123–132.

Dwight W. Read is a Professor at the University of California, Los Angeles. He received his B.A. from Reed College, a M.A. from the University of Wisconsin, and his Ph.D. from UCLA (all in mathematics). He was Co-principal Investigator on the Chevelon Archaeological Research Project and has written extensively on the method and theory of artifact classification. His research interests also include models of hominid origins and the formal analysis of kinship terminology structures. Mailing address: Department of Anthropology, UCLA, Los Angeles, CA 90024.

46. Redman, op. cit. (in note 5); J. Ericson, "Egalitarian Exchange Systems in California: a Preliminary View," in T. Earle and J. Ericson, eds., *Exchange Systems in Prehistory* (Academic Press: New York 1977) 109–126; F. Bove, "Trend Surface Analysis and the Lowland Classic Maya Collapse," *AmAnt* 46 (1981) 93–112.

47. See Read, 1976 op. cit. (in note 5).

48. Compare Redman, op. cit. (in note 5).

49. Compare Nance, 1983 op. cit. (in note 5), which proposes the notion of a Statistical Precision Model versus that of a Discovery Model for sampling.

50. A. Ammerman, A. Voorrips, and D. Gifford, "Toward an Eval-