An Evaluation of Random and Systematic Plot Placement for Estimating Frequency

G.L. WHYSONG AND W.H. MILLER

Abstract

A computer simulation study was conducted to evaluate the effects of pattern on the precision of frequency estimates as determined from random and systematic plot placement. Computer graphics were used to generate artificial population maps containing 40 or 80 clumps of differing spatial intensity with known frequencies of 20, 35, and 50%. The maps were repeatedly sampled both randomly and systematically using a 200-plot sample size to obtain frequency estimates. Three systematic plot spacings (4, 8, and 12) along randomly located transects were evaluated. Analysis indicated that frequency means from systematic plot placement were significantly affected by clumping, pattern intensity, and plot spacing. Random sampling resulted in frequency means that were unaffected by clumping or pattern intensity, and more consistently estimated population frequencies. An evaluation of probabilities of occurrence of Type I errors when statistically comparing frequency estimates from systematic plot placement indicated higher Type I error rates as compared to random sampling.

Key Words: probability, computer simulation, pattern

Field techniques for estimating vegetation frequency for rangeland monitoring and trend evaluation require sampling procedures that will provide reliable data. The integrity of frequency estimates for this purpose is known to be affected by vegetation distribution (Kershaw and Looney 1985, Greig-Smith 1983, West 1985). Finney (1950) found the effect of periodic vegetation pattern to result in less accurate estimates of frequency derived from systematic as opposed to random sampling. Fisser and Van Dyne (1966) reported larger frequencies were obtained from systematic as compared to random sampling of points 3.05 mm in size taken from line intercept transects on foothills grassland in Montana. Schultz et al. (1961) demonstrated that the use of systematically placed points and loops provided overestimates of cover from an artificial community having random dispersion, with cover being analogous to frequency. However, they also presented results indicating that good estimates were obtained from point frame sampling.

Since frequency consists of evaluating presence or absence, computer graphic pixels may be used to indicate the presence of plants by being either lighted or unlighted. Each pixel constitutes a potential sample location within an artificial population. The determination of vegetation frequency in nature is affected by plot size resulting from the spatial interrelationships between vegetation distribution and sample area. Such a restriction does not occur in computer simulation studies as long as the sample area for an individual observation does not encompass multiple pixels. Thus, a pixel may correspond to any plot size and contain the necessary information concerning the presence or absence of vegetation of any size. A similar approach using computer memory locations instead of graphics was employed by Whysong and Brady (1987).

The purpose of this study was to investigate the effect, if any, of aggregated pattern of differing intensity on the reliability of frequency estimates derived from systematic and random sampling. In addition, the effect of plot spacing along transects relative to community dispersion was investigated. Finally, a preliminary evaluation of the applicability of frequency data as determined from random and systematic sampling for use in statistical tests

Authors are associate and assistant professors, respectively, Division of Agriculture, Arizona State University, Tempe 85287.

Manuscript accepted 28 May 1987.

was performed.

Methods

Vegetation Maps

A microcomputer having high resolution graphics capabilities was used to generate simulated vegetation populations of known frequency and pattern. Each population consisted of 128,000 possible sample locations (pixels) arranged on the graphics screen as a rectangle having 200 rows and 640 columns. Vegetation patterns were simulated using randomly lighted pixels (indicating vegetation presence) subject to 3 constraints: (1) The actual specified frequency determined the maximum number of pixels to light, (2) the number of clumps located on the map (40 or 80), and (3) the standard deviation (in pixels) which influenced pattern density.

The center of each clump was randomly located so that no other pixel location within plus or minus 3 horizontal and vertical locations could contain a lighted pixel and the specific pixel located could not have been previously lighted. The number of pixels to be put in a clump was randomly generated to fall within the interval of:

0.5(P/C) and 1.5(P/C)

where:

P = the total number of pixels to light and

C = the total number of clumps.

The last clump located contained the remaining pixels unless there were too many to meet the above constraints. In this case an additional clump was created. Once the center of a clump was determined, the remaining locations to form the clump were randomly located around the center according to the degree of dispersion desired. This dispersion was accomplished by considering the center pixel coordinates as a mean and random numbers were generated to fit a normal distribution with the same mean and a predetermined standard deviation. The algorithm presented by Watson (1981) was used for this purpose. Pixels previously selected were rejected from consideration. Standard deviations of 10, 20, 30, and 40 pixels (coded 4, 3, 2, 1, respectively) were used to regulate pattern intensity. These standard deviations were selected to indicate high and low spatial intensity. The smaller standard deviations resulted in clumps containing more closely spaced (high intensity) patterns while high standard deviations resulted in loosely spaced (low intensity) patterns.

Population maps having actual frequencies of 20, 35, and 50% were generated for sampling purposes. For each actual frequency a total of 8 maps were prepared consisting of all combinations of number of clumps and pattern intensity. Actual frequencies of each map were verified by a separate computer program prior to initialization of sampling. A 20% frequency map generated with 80 clumps and a standard deviation of 20 pixels is presented in Figure

Map Sampling

Each of the 24 maps was sampled both systematically and randomly by the computer. Systematic sampling consisted of randomly locating 4 starting locations and checking for presence or absence at 50 pixel locations along a transect oriented along the longest axis of the rectangular map. Each sample consisted of 200 observations. A 200-plot sample size was selected since it is similar to that frequently used for rangeland monitoring. Plot spacing was

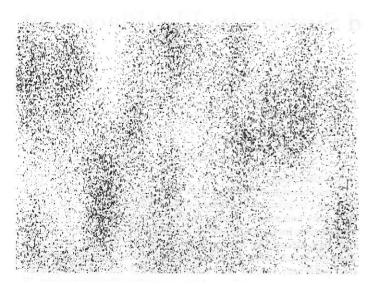


Fig. 1. Computer generated vegetation map having a population frequency of 20%, 80 clumps, and pattern intensity determined by using a standard deviation of 20 (pixels) within each clump.

evaluated by using 4, 8, and 12 pixel spacings between observations along transect lines. Random starting locations were determined such that all possible pixels in the population were potentially accessible. In order to obtain sufficient precision of parameter estimates and data amenable for statistical analysis, maps were independently sampled to provide 5 replications of each frequency, clump number, pattern intensity and spacing combination. Each replication consisted of the mean from 100 samples of 4 transects. Thus, each of the 72 possible treatment means was determined from 2,000 randomly located transects (100,000 plots). Care was taken to assure that the random number generator was periodically reseeded to avoid repetitive random number sequences.

Random sampling was conducted using the same number of replications and plots per treatment as systematic sampling. Observations were made at random from vegetation maps with the only restriction that a location previously measured during the same sample of 200 observations could not be evaluated again.

Since only 200 observations were made per sample from a population consisting of 128,000 possible measurable locations, no finite correction was necessary. Cochran (1963) and Scheaffer et al. (1979) indicated that no finite correction is required when sample size does not exceed 5% of the population. In this study, the computer evaluated less than 1% of the population when conducting each independent sample.

Data Analysis

Frequency data commonly fit the binomial distribution. However, means from non-normal distributions as determined from independent samples will tend toward normality according to the central limit theorem (Zar 1984). In addition, estimates from a population containing an actual frequency of 50% approximate the normal distribution. These properties justify the use of analysis of variance (ANOVA) procedures to compare sample results. Frequency means as estimated from systematic sampling were analyzed separately for each population frequency. A 3 factor factorial was used in each analysis to evaluate the effects of clumps, pattern intensity, and plot spacing. Two factor factorial ANOVA was used to determine potential differences due to clumps and pattern intensity for each population frequency sampled randomly.

Following evaluation of the ANOVA results, systematic and random sampling were investigated further using the population consisting of 20% frequency and 80 clumps. A total of 20,000 independent samples were taken randomly and systematically

from the community by the computer. Systematic sampling was conducted separately using each plot spacing. The first phase of the evaluation was to determine the probability that the sample estimates would fall within the 95% binomial confidence interval of the actual map frequency. The second phase was designed to compare the reliability of sample estimates as determined from random and systematic sampling for use in statistical tests for detecting differences. Greig-Smith (1983) and Hironaka (1985) suggest the use of Chi-Square procedures for such tests. Zar (1984) points out that the two-tailed test of proportions is equivalent to Chi-Square and the 2 are interchangeable. The 0.05 probability level was used for a two-tailed test of proportions to evaluate the probablity of committing a Type I error when the same population was repeatedly measured using random and systematic sampling. Each probability was estimated using 19,999 tests by comparing the prior frequency value to the current sample estimate as the computer conducted the 20,000 samples. A correction for continuity was not employed as preliminary simulations indicated that correction resulted in overly conservative tests. Whysong and Brady (1987) found the correction for continuity too conservative when using the same test for evaluating the effect of sample size on probabilities of detecting frequency differences.

Results and Discussion

Table 1 presents F values as determined by ANOVA and their statistical significance for each frequency, factor and type of sampling. Results from random sampling indicated nonsignificance among number of clumps, pattern intensity, or the interaction of these 2 factors at all 3 population frequencies studied. These results

Table 1. Calculated F values from analysis of variance of systematic and random sampling results.

Source	20% Freq.		35% Freq.		50% Freq.	
	Syst.	Rand.	Syst.	Rand.	Syst.	Rand.
Clumps (C)	0.62ns	1.33ns	0.68ns	0.08ns	9.85**	0.13ns
Intensity (I)	17.54**	0.17ns	7.62**	1.14ns	6.18**	1.18ns
Linear	6.47*		10.01**		12.48**	
Quadratic	34.60**		4.87*		0.09ns	
Remainder	18.01**		7.98**		5.96*	
Spacing (S)	20.90**		8.60**		6.30**	
Linear	20.59**		13.29**		6.45*	
Remainder	21.21**		3.91ns		6.15*	
$C \times I$	23.94**	1.18ns	21.71**	2.80ns	10.74**	0.38ns
Clin. \times Ilin.	3.11ns		11.88**		2.06ns	
Clin. × Iquad.	2.79ns		38.65**		29.06**	
Clin. X Irem.	65.23**		14.64**		1.11ns	
$C \times S$	6.58**		1.22ns		0.19ns	
Clin. × Slin.	5.19*				B. G.	
Clin. × Srem.	7.97**					
$I \times S$	2.04ns		1.13ns		2.23ns	
$C \times I \times S$	0.49ns		2.81*		6.33**	

Significant F value (P<0.05).

ns Non-significant F value (P>0.05).

indicate that random sampling estimated the same mean frequency for all 8 population maps within each actual frequency. Consequently, it would appear that frequency means as determined by repeated random sampling are unaffected by the number of clumps and the intensity of pattern within clumps.

Numerous highly significant results were found when systematic sampling was employed at each actual frequency (Table 1). The number of clumps was significant only in the 50% population but cannot be de-emphasized due to the significant interaction between clumps and pattern intensity that occurred in all 3 actual frequencies. Pattern intensity was highly significant in every case, with response surface evaluation indicating the effect was strongly nonlinear, beyond the quadratic. Systematic plot spacing was also

^{**} Significant F value (P<0.01).

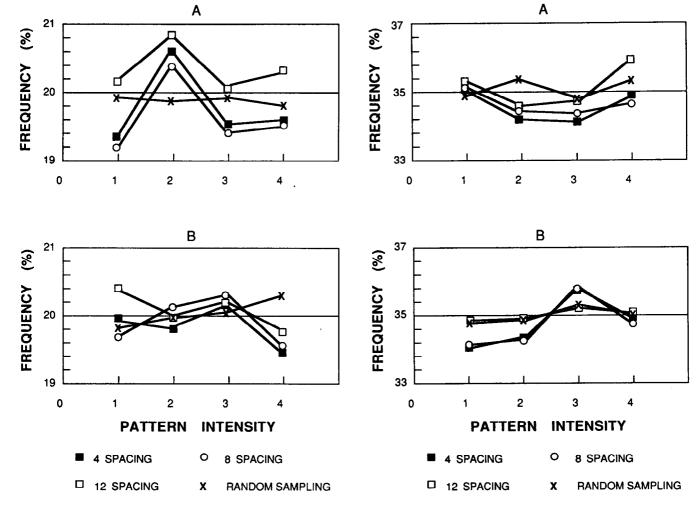


Fig. 2. Frequency means from the 20% frequency population using systematic plot placement and random sampling at 4 intensities of pattern with (A) 40 clumps and (B) 80 clumps.

Fig. 3. Frequency means from the 35% frequency population using systematic plot placement and random sampling at 4 intensities of pattern with (A) 40 clumps and (B) 80 clumps.

significant in all ANOVA's and demonstrated nonlinearity in the 20% and 50% frequency populations. In addition, a nonlinear interaction between clumps and plot spacing occurred at 20% frequency. These results indicate that problems may exist concerning the reliability of frequency means as determined from systematic sampling.

Analysis of variance indicated, within the 8 maps systematically sampled at each actual frequency, that populations having the same frequency were significantly different. These differences were the result of 2 major factors. First, the spatial arrangement (pattern) seems to influence the frequency estimate in a nonlinear manner. As pattern changes, particularly intensity, the sample mean will also change. It is possible that such an effect is independent of sample size and will manifest itself regardless of the number of transects measured. Second, the method of employing the systematic sample will affect the frequency estimate. The spacing between plots resulted in significantly different frequency means. In addition, plot spacing was involved in a significant interaction with pattern factors at each frequency studied. These results suggest that frequency means from different systematic plot spacings may not be comparable.

The effects of plot spacing and intensity of pattern are illustrated for the 20% frequency populations in Figure 2. At 40 clumps per population, frequency means demonstrated an oscillatory pattern around the true frequency (Fig. 2a). Plot spacing means appear fairly consistent relative to each other; however close inspection suggests each spacing may be operating cyclically, but at a slightly

different phase. the means from random sampling show greater consistency. A different cyclic response appears when the number of clumps is doubled to 80 (Fig. 2b). Once again, random sampling provided more consistent mean estimates of frequency at different pattern intensities than any single systematic plot spacing, although differences were smaller than at 40 clumps.

Cyclic behavior of systematic frequency means was also evident relevant to intensity of pattern at 35% frequency (Fig. 3). At 40 clumps, spacing results appeared to behave similar at lower pattern intensities with divergence occurring as intensity increased (Fig. 3a). Both the 4 and 8 spacing tended to respond similarly to pattern intensity at 80 clumps while the 12 spacing showed minor oscillations around the actual frequency (Fig. 3b). It is interesting to note that frequency means from the systematic 12 spacing appeared as accurate as those obtained from random sampling. However, the trend of the 12 spacing matches that of the 4 and 8 spacings. Only the magnitude of the oscillations are different.

Figure 4 presents the spacing results relative to pattern intensity for the 50% population maps. Cyclic responses of plot spacing in relation to clumping and intensity of pattern appear evident. Closer association tended to occur between the 4 and 8 spacings whereas the 12 spacing reacted somewhat differently possibly due to differences in cycle amplitude.

Kershaw and Looney (1985) and Creig-Smith (1983) discussed the effect of spatial periodicites due to vegetation pattern on estimates derived from systematic samples. They both point out that regularly spaced observations may lead to biased sample results if

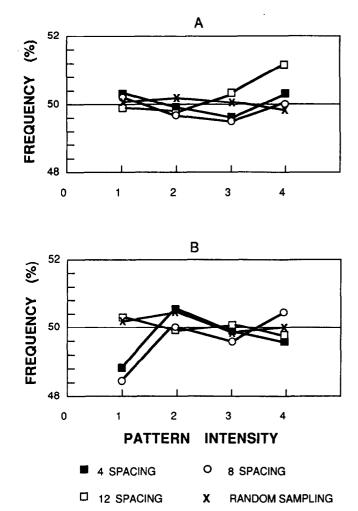


Fig. 4. Frequency means from the 50% frequency population using systematic plot placement and random sampling at 4 intensities of pattern with (A) 40 clumps and (B) 80 clumps.

spacing corresponds to topographical irregularities which affect vegetation distribution. Greig-Smith (1983) further indicated that frequency was not independent of plant density and pattern. The effect of inherent periodic or cyclic variation in populations can affect results obtained from systematic samples (Cochran 1963). Additionally, size and arrangement of vegetation clumps plus the spatial distribution of plants within clumps may easily result in periodicities of plant abundance occurring within vegetation communities.

The computer generated maps used in this study did not incorporate topographical features, nor was any attempt made to simulate topographic effects on population distribution. Clumps were randomly located subject to the restrictions outlined earlier. Intensity of pattern (spatial arrangement within clumps) was controlled by the standard deviation selected when a clump was formed. Thus, each map series consisted of a gradient of high intensity (closely spaced individuals) to low intensity (widely spaced individuals). As intensity decreased or the number of clumps being formed increased from 40 to 80, the frequency of clumps being formed in close proximity to each other increased. Under these conditions several clumps could merge together forming a population containing clumps of different sizes. The map shown in Figure 1 illustrates this aspect. Since the exact nature of each clump was not maintained, an assessment of periodicity relative to the results obtained from systematic plot placement is difficult, if not impossible. However, the data presented in Figures 2 through 4 suggest a strong cyclic behavior dependent on the interaction of clumping,

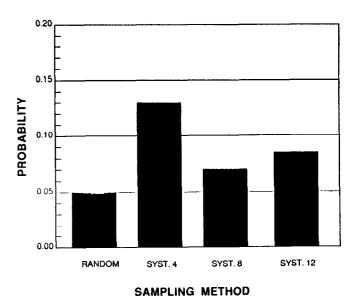


Fig. 5. Probability of frequency estimates from random and systematic plot placement occurring within the P=0.05 binomial confidence limits (N=20,000).

spatial intensity, and plot spacing. The inability of sample method to control the accuracy of frequency sample means is of primary importance and raises doubts as to the applicability of systematic plot placement for frequency determinations.

An inspection of the means presented in Figures 2 through 4 may cause one to question the previous discussion. The greatest deviation of a systematic treatment mean from the population parameter was about 1.4% (Fig. 4b). Most means were appreciably closer. However, due to the large number of observations contributing to each treatment mean (100,000) they have been estimated with a higher degree of precision than field sampling would allow.

Figure 5 illustrates the results obtained from comparisons of frequency estimates from random and systematic sampling to binomial confidence limits. These data were obtained by sampling the same map (Fig. 1). Estimates from repeated random samples should fall within the 95% binomial confidence limits 95% of the time and correspond to the probability of detecting no difference in the population. Conversely, 5% (P=0.05) of the estimates should occur outside the confidence region and correspond to the probability of a Type I error occurring in a statistical test. As can be seen, the 0.05 probability was accurately maintained when estimates obtained by random sampling were evaluated against the 95% confidence limits.

Cochran (1963) indicated that a systematic sample may be considered as a simple random sample with N=1. However, when periodic variation is present, unbiased estimates of parameters are unlikely. Mean frequencies as determined from the 4, 8, and 12 spacings were calculated as the mean of 4 randomly located transects of 50 plots each. Thus, 4 independent frequency determinations comprised each systematic mean. The results of comparing these means to the 95% binomial confidence limits are shown in Figure 5. In each case, probabilities obtained exceeded the 0.05 level.

Particularly noticeable is the 4 spacing which exceeded the random sampling results to the greatest degree. These results suggest that frequency means derived from systematic plot placement along randomly located transects are unlikely to fit the binomial distribution. Greig-Smith (1983) indicated that frequency derived from a single systematic sample does not fit the binomial distribution.

A two-tailed test of proportions was used to compare frequency means from systematic samples at each plot spacing. Measurements were taken from the same population map as previous

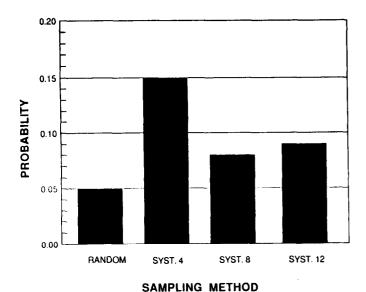


Fig. 6. Probability of a Type I error (P=0.05) from statistical comparisons of frequency estimates using random and systematic plot placement (N=19,999).

comparisons. The test of proportions provides identical results to those obtained from Chi-Square contingency tables usually recommended for frequency comparisons Zar (1984). Results indicated that the probability of Type I errors occurring were greater than the 0.05 probability expected at all systematic plot spacings studied (Fig. 6). However, the results obtained from random sampling agreed with the expected 0.05 probability. Since the probability of committing a Type I error varied due to systematic plot placement and was of greater magnitude then expected, the suitability of these tests for detecting changes in frequency determined from systematic plots appears questionable. Since the probability estimates were determined from only 1 population map, no information was obtained concerning the possible effects of other intensities of pattern and spacing combinations on the occurrence of Type I errors.

Conclusions

The results obtained from this study suggest that the use of systematic sampling along randomly located transects for assessing vegetation frequency should be questioned. The distribution of

plant species rarely appears either random or uniform in natural communities. Our data suggest that frequency obtained from systematic observations of populations demonstrating varying degrees of aggregation may be biased. Furthermore, the amount of bias will not only result from the systematic spacing selected by the investigator, but may be influenced by the distribution of vegetation within the community being sampled. Thus, little control of bias is possible by the investigator as long as systematic sampling for frequency is employed. The severity of bias is sufficient to result in large increases in the probability of committing Type I errors when traditional statistical tests are employed to detect differences in frequency. In light of these results, and the increasing interest in frequency for evaluation of range trend, additional research may be necessary to substantiate and further clarify the effects of vegetation pattern and plot spacing on frequency estimates.

Literature Cited

Cochran, William G. 1963. Sampling techniques. John Wiley and Sons, Inc., New York.

Finney, D.J. 1950. An example of periodic variation in forest sampling. Forestry 23:96-111.

Fisser, H.G., and G.M. Van Dyne. 1966. Influence of number and spacing of points on accuracy and precision of basal cover measurements. J. Range Manage. 19:205-211.

Greig-Smith, P. 1983. Quantitative plant ecology. 3rd. ed. Univ. of California Press. Berkeley.

Hironaka, M. 1985. Frequency approaches to monitor rangeland vegetation. Symp. on use of frequency and density for rangeland monitoring. William C. Krueger, Chairman. Proc. 38th Annual Meeting, Soc. for Range Manage. Feb. 1985. Salt Lake City. Soc. for Range Manage. 84-86.

Kershaw, K., and J.H. Looney. 1985. Quantitative and dynamic plant ecology. 3rd ed. E. Arnold, Publ. Ltd. Baltimore.

Scheaffer, Richard L., William Mendenhall, and Lyman Ott. 1979. Elementary survey sampling. 2nd. ed. Duxbury Press, North Scituate, Mass.

Schultz, A.M., R.P. Gibbens, and L. DeBano. 1961. Artificial populations for teaching and testing range techniques. J. Range Manage. 41:236-242. Watson, Hugh J. 1981. Computer simulation in business. John Wiley &

Sons, Inc. New York.

West, N.E. 1985. Shortcomings of plant frequency-based methods for range condition and trend. William C. Krueger, Chairman. Proc. 38th Annual Meeting, Soc. for Range Manage. Feb. 1985. Salt Lake City. Soc. for Range Manage. 87-90.

Whysong, G.L., and W.W. Brady. 1987. Frequency sampling and Type II errors. J. Range Manage. 40:472-474.

Zar, Jerrold H. 1984. Biostatistical analysis. 2nd ed. Prentice-Hall Inc., Englewood Cliffs, N.J.