STABILIZATION OF THE MEAN AS A DEMONSTRATION OF SAMPLE ADEQUACY

by

David L. Clark

Abstract Adequate sampling of plant populations can be problematic. Commonly estimated population parameters are seldom normally distributed, and minimum sample sizes that are based on confidence intervals for normal populations may overestimate how many samples are required to accurately estimate production and density, and sometimes cover. Alternative procedures for determining sample adequacy based on the ideas underlying the species-area curve have been proposed, but few studies have evaluated the effectiveness of these procedures on field data. Running mean calculations from undisturbed and reclaimed vegetation communities in New Mexico indicate that a stable estimate of the mean is often obtained after 30 to 40 production and density samples. Mean cover estimates typically stabilize after 15 to 20 samples. For data where the mean fails to stabilize after 40 samples, quadrat size problems are indicated. A case is made for the adoption of a minimum sample size of 30 and a maximum sample size of 40, and calculation of sample adequacy using the standard deviation of the consecutive means, instead of the standard deviation of the individual samples. The merits of appropriately-sized quadrats when measuring plant production and density are also illustrated.

Additional Key Words: mined land, reclamation, vegetation inventory

Introduction

Problems with demonstrating adequate estimation of vegetation parameters have been recognized for many years. A principle difficulty stems from attempting to use minimum sample size formulas that are based on confidence intervals for normally distributed populations. Vegetation parameters such as production and density seldom appear to be normally distributed. Due to the clumped dispersion of vegetation and the generally small plot sizes used for sampling, these parameters typically are best described by the negative binomial distribution. The distribution of vegetation cover may be more nearly normal, but the binomial model may be the best fit for cover. Using the equations derived from the normal model for vegetation populations having clustered dispersion and binomial or negative binomial distributions often results in large minimum sample size estimates that remain large even when additional samples are collected. Since the populations are not randomly dispersed, small sample plots occasionally fall into either voids or dense patches. As the number of samples increases, the variance of the individual samples does not decline, but the sample mean does stabilize and the variance between successive estimates of the mean does decline. This result is consistent with the Central Limit Theorem, which predicts that the successive means will be normally distributed.

Graphical approaches based on the concept of the species-area curve (Cain 1938) for the demonstration of sample adequacy for vegetation studies have been suggested in several methodology texts and manuals (Pieper 1978, Grieg-Smith 1983, Bonham 1989, Krebs 1989). A specific recommendation of this approach for vegetation analyses on mine lands was made by Sowards (1983). The idea is to plot the running means of the measured vegetation attribute on the y-axis, and the number of samples from which each successive mean was calculated on the x-axis. Stabilization of the running mean occurs as sample size increases, and the running mean attains a kind of inertia, or resistance to the influence of extreme values. This stabilization is plainly illustrated by the decreasing influence of extreme shrub density values

[1]Paper presented at the 2001 National Meeting of the American Society for Surface Mining and Reclamation, Albuquerque, New Mexico, June 3-7, 2001. Pub. by ASSMR, 3134 Montavesta Rd., Lexington, KY 40502.

[2]David L. Clark is Surface Mining Reclamation Specialist, Mining and Minerals Division, New Mexico Energy, Minerals and Natural Resources Department, 1220 South Saint Francis Drive, Santa Fe, NM 87505.

Proceedings America Society of Mining and Reclamation, 2001 pp 65-69
DOI: 10.21000/JASMR01010065

https://doi.org/10.21000/JASMR01010065

in Figure 1. When successive means remain within a specified limit of variability over a specified sample interval, the population parameter could be considered adequately estimated. Sowards (1983) proposed a variability limit of ± 2.5% over 10 consecutive calculations of the running mean.

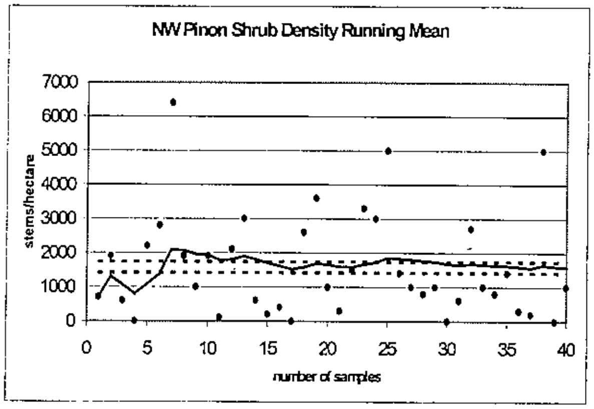

Figure 1. Individual quadrat data and running mean estimate of shrub density for a reclaimed site. The dotted lines depict ± 10% of the mean estimate after 40 samples were taken.

Artificial random populations were used by Schultz and others (1961) as an aid for teaching range cover measurement techniques. Successive sample means stabilized to within ± 10% of the true population mean after 30 to 40 transects were read using line intercept, line point, or point intercept measurement methods. Whether similar results could be obtained when sampling real, non-randomly dispersed vegetation was not determined.

Objectives

The stabilized-mean approach for demonstrating sample adequacy has not been widely adopted by either the scientific community or regulatory agencies. A lack of axis-scale standardization and the distortion that occurs when the x/y axis ratios are varied (Cain 1938) have probably contributed to this indifference. Also, a measure of the confidence level or repeatability of the mean estimate is not provided by the graphical method, and few studies exploring the validity of the method for vegetation studies have been presented. Large data sets have recently become available from coal mine bond release evaluations conducted in New Mexico, and were used in this paper to explore the efficacy of the stabilized-mean approach. The effects of increasing plot size on density and production estimates were also investigated.

Methods

For each of the studies evaluated in this paper, consistent methods were used to collect cover and production data. Gentle slopes, commonly 1-2% and always <10%, were sampled on both reclaimed and undisturbed areas. Cover was read at 0.5-m intervals on 50-m long, randomly located point intercept transects. Above ground plant biomass produced during the current growing season was harvested from 1x1-m quadrats placed at the starting point of each shrub density transect. Shrub stocking was estimated by counting shrubs and subshrubs rooted within either 50x2-m or 50x4-m belt transects, depending on the study.

Minimum sample sizes were calculated using the Cochran (1977) formula:

$$n_{min} = \frac{(ts)^2}{(0.10\bar{x})^2} \qquad (1)$$

where

- t is the tabular t value for a preliminary sample with n-1 degrees of freedom and a two-tailed significance level of $\alpha = 0.10$
- s is the standard deviation of a preliminary sample
- $\bar{x}$ is the sample mean of a preliminary sample.

Results and Discussion

Successive means for shrub density data from a reclaimed site in northwest New Mexico are depicted in Figure 1. The dotted lines represent ± 10% of the final estimate of mean density (1582 stems/hectare). After taking 40 samples, the formula (1) result for the minimum sample size required to estimate the mean with a precision of ± 10% was 262. It does not seem reasonable that an additional 222 samples are needed to achieve the desired level of precision, since a reduction in the influence of the extreme data values is seen as sample size increases. Similar patterns are displayed by the running means for cover (Figure 2) and shrub density (Figure 3) from undisturbed plant communities.

Forty samples may be insufficient to determine if the stability attained by the running mean is transitory. A sample size of 120 was used to assess reclamation shrub stocking at a mine in west-central New Mexico (Figure 4). The running mean stabilized at about 1430 stems/hectare between 30 and 40 samples, but then stabilized at about 1300 stems/hectare when 50 to 60 samples were obtained. A lower, final estimate of about 1190 stems/hectare was provided when 80 to

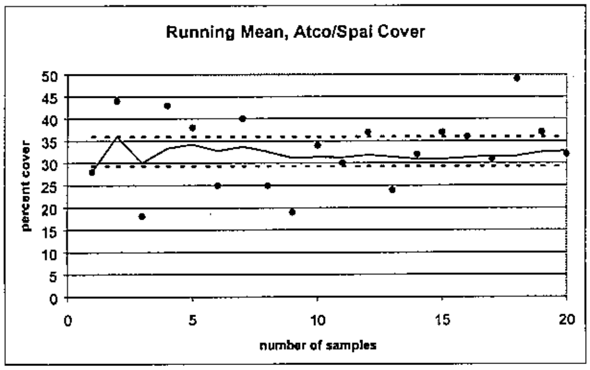

Figure 2. Individual transect data and running mean estimate of cover for a native shrub community. The dotted lines depict $\pm$ 10% of the mean estimate after 20 samples were taken. The minimum sample size calculated after taking 20 samples was 19.

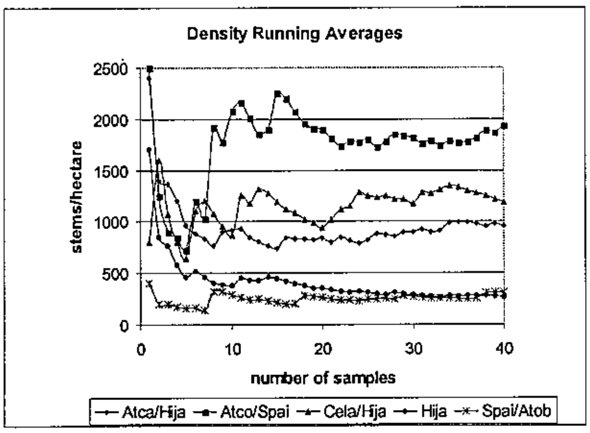

Figure 3. Running mean estimates of shrub density for 5 undisturbed vegetation communities.

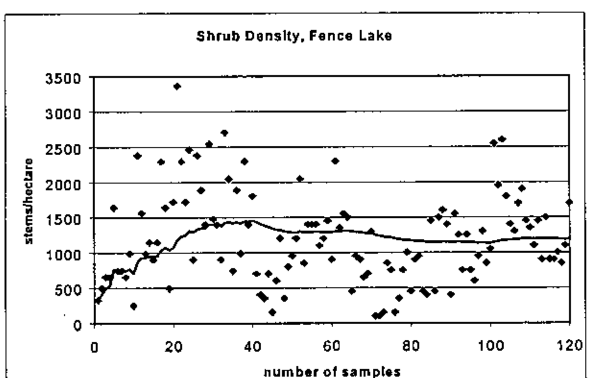

Figure 4. Individual belt transect data and running mean estimate of shrub density for a reclaimed site. The minimum sample size calculated after taking 120 samples was 81.

120 samples were taken. For these data, the sample adequacy formula indicated that the mean was estimated to within 20% of the true value, with 90% confidence, after 40 samples. A minimum sample size of 81 was calculated by formula (1). That sample size conforms very well to the point at which the running mean stabilized at virtually the same value as the final estimate. Stabilization of 10 consecutive running means to within $\pm$ 2.5% of the final estimate occurred after 112 samples. Stabilization to within $\pm$ 10% of the final estimate occurred after 57 density samples.

Figure 4 might initially lead one to conclude that stabilization of the mean is more of an illusion than a meaningful or useful tool. It appears that regions of more and less variable shrub stocking were present on the reclamation, and as sampling progressed from one region to another, the value of the running mean shifted. It might be instructive to stratify the reclamation into regions of different variability, and either conduct two-stage sampling or separately evaluate the sample size for each region. However, a simpler approach is to determine whether increasing the plot size would reduce the variability of the stocking estimate. Figure 5 depicts the density estimates and running mean that resulted when the original 120 samples were divided into 40 sets of 3 randomly selected belt transects each. The stocking estimates from the 3 transects were added to create 40 samples. The virtual effect was to treble the size of each belt transect. The mean stabilized after 25 of the enlarged transects. A minimum sample of 43 enlarged transects was calculated by formula (1). The apparent inadequacy of 40 samples that is observed in Figure 4 is thus more a function of less-than-optimal plot size, rather than sample size.

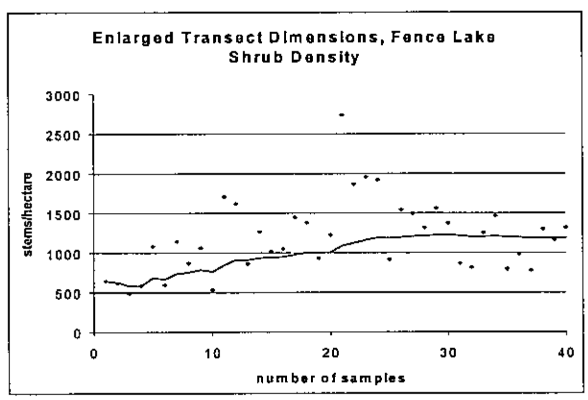

Figure 5. Running mean estimate of shrub density for a reclaimed site after trebling the plot size. The minimum sample size calculated after taking 40 samples was 43.

Production data are seldom described as normally distributed. Figure 6 depicts the results of clipping 200 1x1-m production plots on a reclaimed area. Stabilization of 10 consecutive running means to within $\pm$ 2.5% of the final estimate occurred after 151 samples, and stabilization to within $\pm$ 10% of the final estimate occurred after 108 production samples. The minimum sample size calculated by formula (1) for these data was 197 clip plots.

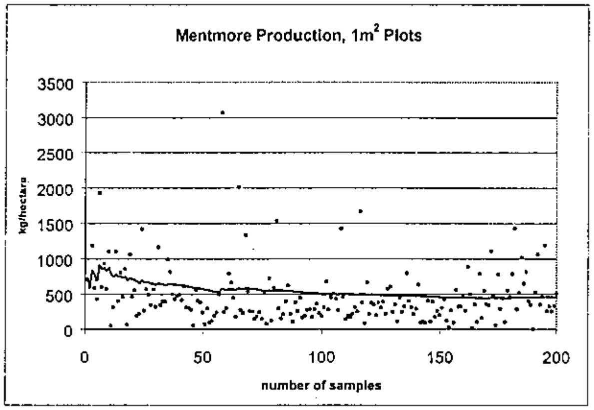

Figure 6. Individual quadrat data and running mean estimate of annual production for a reclaimed site made by clipping $1m^2$ plots. The minimum sample size calculated after taking 200 samples was 197.

The 200 samples were divided into 40 sets of 5 randomly selected plots, and the 5 plots were added to create 40 samples of $5m^2$ each (Figure 7). Increasing the plot dimensions resulted in a calculated minimum sample size of 31 enlarged clip plots. The running mean stabilized after 5 enlarged clip plots were measured.

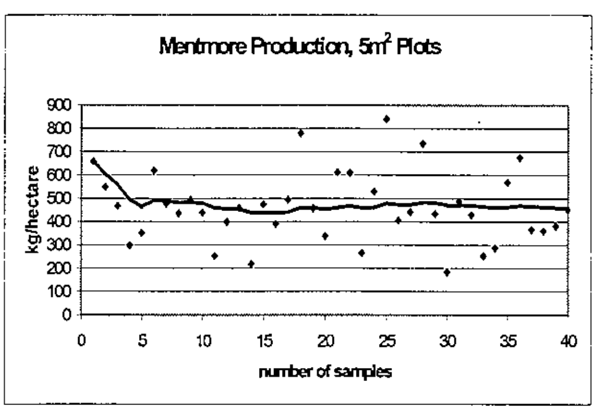

Figure 7. Running mean estimate of annual production for a reclaimed site made by clipping $5m^2$ plots. The minimum sample size calculated after taking 40 samples was 31.

The efficiency gained by clipping larger plots is impressive, and illustrates the importance of tailoring quadrat size to the type of vegetation that is being sampled.

Conclusions and Recommendations

The large data sets examined in this paper revealed both strengths and weaknesses with graphing running means as a demonstration of sample adequacy. With the graphical approach we see a normalization of consecutively calculated means due to the Central Limit Theorem. Unfortunately, the graphical approach as proposed in most texts and manuals does not provide a confidence level or other measure of the repeatability of the mean estimate. Stabilization of the mean to within a reasonably tight variability limit, e.g., Sowards' (1983) proposal of $\pm$ 2.5%, may require nearly as many or (as in the shrub density example featured in this paper) even more samples than calculated by formula (1). Running means may stabilize at more than one value (Figure 4) or gradually approached a final estimate without a clear point of stabilization (Figure 6).

The running mean approach does support the idea that 30 to 40 samples should be sufficient to provide reliable parameter estimates, if the data are normal, or can be normalized (Figures 5 and 7). This leads one to consider whether it would be legitimate to take 30 to 40 samples, calculate the running means, and use the standard deviation of the running means in formula (1) to determine if an adequate sample has been obtained. Used in this manner, running mean calculations would be akin to normalizing the data by transformation. Data transformation is routinely accepted by both the scientific and regulatory communities, and calculated means are certainly more intuitive and interpretable than calculated logarithms, reciprocals, or arcsines.

Applying the procedure proposed above to the shrub density data reviewed in this paper, the minimum sample size needed to obtain 90% confidence that the running mean is within 10% of the true mean, calculated from the first 40 samples, is 15. Similarly, due to normalization by the Central Limit Theorem, we can predict that sample adequacy will be demonstrated after taking 30 to 40 samples for virtually any vegetation parameters, if the standard deviations of the running mean are used in formula (1). There may be cases in which extremely variable data will require more than 40 samples. An adequacy calculation by the proposed method would identify those cases and alert the investigator that there may be problems with sampling design, such as improper stratification or quadrat size.

The notion that vegetation parameters are not normally distributed can often be attributed to using small plots to sample sparsely and contagiously dispersed plants. There are probably scales at which all vegetation parameters are normally distributed, and more attention should be given to the identification of those scales. Unfortunately, investigators often simply rely on standard plot dimensions and hope for the best. Quadrat dimensions selected to ensure that each sample captures at least a few stems (or grams of annual production), but not too many, will minimize the sample variance and the required sample size.

Using the standard deviation of the running means in formula (1) addresses the concerns that have been expressed over the years about adequate sample size for vegetation inventories. Intuitively reasonable sample sizes, a calculation of the confidence we have in our parameter estimate, and an indication of poor sample design can all be obtained.

Acknowledgements

The data presented in this paper were provided by the following New Mexico coal mine operators: San Juan Coal Company, Carbon Coal Company, and the Salt River Project. Their cooperation in publicizing this information is appreciated.

Literature Cited

Bonham, C. D. 1989. Measurements for terrestrial vegetation. John Wiley and Sons, New York.

Cain, S. A. 1938. The species-area curve. Amer. Midl. Nat. 15:529-566.
https://doi.org/10.2307/2419922

Cochran, W. G. 1977. Sampling techniques. John Wiley and Sons, New York.

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

Krebs, C. J. 1989. Ecological methodology. Harper and Row, New York.

Pieper, R. D. 1978. Measurement techniques for herbaceous and shrubby vegetation. Dept. of Animal and Range Sciences, New Mexico State University, Las Cruces.

Schultz, A. M., R. P. Gibbens, and L. DeBano. 1961. Artificial populations for teaching and testing range techniques. J. Range Management 14:236-242.
https://doi.org/10.2307/3894738

Sowards, W. E. 1983. Achieving an adequate sample size for vegetation analyses. Pages 63-71 in E.F. Redente, W. E. Sowards, D. G. Steward, and T. L. Ruiter (ed.). Western coal mining regulatory issues: land use revegetation and management. Proc. Symp. West. Reclam. Grp. August 1983. Colorado St. Univ. Range Science Series 35.