An evaluation of sample adequacy for point analysis of ground cover

L. HOFMANN AND R.E. RIES

Abstract

A sample adequacy equation is recommended to calculate the number of point frame samples required for measuring cover on reclaimed land and to assure sample size is adequate to meet Federal rules and regulations. When applied to field data, mining companies and regulatory agencies often find the sample number requirements estimated by the equation are so large that the equation is impractical to use. This equation was studied by sampling with 20, 40, 60, and 100 sets of 10-point frames on 12×67-m areas of grazed and ungrazed mixed prairie at Mandan, N.D., and by examining cover data collected with 60 sets of frames from 1.86-ha pastures on reclaimed mined land and native range near Center, N.D. Increasing the number of frames used did not produce more consistent mean values nor did variance decrease. Both total ground cover and bare soil ground cover measure the amount of cover protecting the soil from soil loss; yet, the formula estimated that 1 frame was required to measure total cover whereas 10,086 frames were required to measure bare soil within 10% of the mean with 90% statistical confidence. Histograms suggest cover components that comprise a small percentage of the total ground area are distributed in a Poisson rather than normal fashion; therefore, the equation does not provide a good guide for determining how many samples are required. Another equation, appropriate for binomial and Poisson variates is suggested as a solution to the problem.

Key Words: sample size, formula, reclaimed mined land, native prairie, data transformation, proportion or percentage data

The point quadrat method (Levy and Madden 1933) has been extensively used to determine canopy or basal ground cover of native or introduced vegetation in northern Great Plains grassland. Point analysis also is recommended for measuring cover on reclaimed surface-mined land to assess successful revegetation. Federal rules and regulations (Federal Register Vol. 44, No. 50, 13 March 1979) set statistical requirements for revegetation of surface-mined land as follows: "the ground cover and productivity of the revegetated area shall be considered equal if they are at least 90% of the ground cover and productivity of the reference area with 90% statistical confidence". The equation

$$n = \frac{t^2 s^2}{d^2} \qquad (1)$$

where n = number of samples required, t = t-distribution value for a given level of confidence, s^2 = the estimate of the variance from sampling, and d = level of accuracy desired for the estimate of the mean, has been suggested by Bonham (1982), Chambers and Brown (1983), and others to calculate how many samples must be taken to meet federal requirements. Equation (1) is based on the assumption that the data are normally distributed.

When mining companies applied equation (1) to point frame data, the calculated number of frames required was often impractical to attain because numbers were so large. The large sample requirement also became a concern for regulatory agencies that were attempting to develop guidelines to meet the law.

Sowards (1983) reviewed some of the difficulties mining companies had with using the adequacy formula for determining sampling size required to estimate vegetation production. He offered an alternate nonstatistical approach of graphing the running consecutive means as a function of sample size to evaluate sample adequacy.

Earlier studies have been conducted to determine the number of points required to adequately sample perennial grass cover. Levy (1933) suggested that 100 points were sufficient for charting dominants within a pasture sward, but 400 to 500 points were required to chart lesser species. Drew (1944) compared readings of 10, 20, and 30 needles in 0.5-m quadrats and reported estimates of total hits from 10 needles were similar to estimates from 30 needles. Crocker and Tiver (1948) reported reliable vegetation analysis for survey work was obtained by taking 300 to 500 points per field or pasture ranging in size from 5.7 to 60.7 ha.

Whitman and Siggeirsson (1954) compared line intercept, all-contact points, and basal-contact points for analyzing mixed-grass rangeland vegetation in western North Dakota. They calculated the numbers of observations required to produce sampling errors of 5% and 10% of the means using the equation

$$N = \frac{S^2}{(Sx)^2}$$

where S = standard deviation derived from the plot data and Sx = 5% and 10% of the mean density of a species or group. They suggested a minimum working basis of 23 line transects, 1,400 all-contacts points, or 3,600 basal-contact points would be required for survey work in the type and density of vegetation they sampled.

The purpose of this study was to determine how the number of point frame samples affects the estimated mean and variance of ground cover components and to examine how sample size estimates calculated by the equation (1) change as sample size changes.

Methods

Point frame data were obtained from 2 areas: native mixed-prairie pasture near Mandan, N.D., during August, 1986, and from reclaimed mined land and native pasture near Center, N.D., (Hofmann and Ries 1989) during May, 1985. Cover was estimated using a standard frame with 10 pins spaced approximately 5.0 cm apart. The frame was held vertical, pin contact with above-ground vegetation was ignored, and soil surface contacts (basal hits) were counted.

At Mandan, live and litter basal ground cover, and bare ground from 12×67 m plots were repeatedly sampled using 20, 40, 60, and 100 sets of frames. Each set of data, i.e., 20, 40, 60, or 100 frames, was obtained separately with frames placed at random locations over each plot area. One plot was a moderately grazed (1.2 AUM ha^{-1}) pasture and the other on an adjacent ungrazed exclosure. Dominant species were western wheatgrass [*Pascopyrum smithii* (Rydb.) Löve], green needlegrass (*Stipa viridula* Trin.), needle-and-thread (*S. comata* Trin. and Rupr.), blue grama [*Bouteloua gracilis* (H.B.K.) Lag.], and sedges (*Carex* spp.).

545

Reclaimed-land vegetation consisted primarily of smooth brome-grass (*Bromus inermis* Leyss.), crested wheatgrass [*Agropyron desertorum* (Fisch. ex. Link) Schultes], intermediate wheatgrass [*Thinopyrum intermedium* (Host) Barkworth & D.R. Dewey subsp.], and alfalfa (*Medicago sativa* L.). Adjacent native species were similar to those at the Mandan site. Five frames were read at random within 6 general locations across the 1.86-ha grazed area of 2 replicated reclaimed and native pastures for a total of 60 frames.

The equation

$$\hat{n} = \frac{t^2 s^2}{(\bar{x} * .10)^2} \qquad (2)$$

where $\bar{x}$ = the sample average was used to calculate the number of frames required to adequately sample basal cover within 10% of the population mean at the 90% confidence level required by Federal regulations. One-tailed t values were used. Cover means were obtained by multiplying pin contacts by 10 to provide cover values that total 100. Means, variance, and numbers of point frame samples required are reported for litter, live, total (litter + live), and bare (100 − total) cover for all areas. Individual major species were estimated from the reclaimed and adjacent native pasture. Histograms of the sampling distributions for each of the various cover measurements were plotted using SAS/GRAPH procedures (SAS 1985) to visually assess normality.

A square root transformation, generally used to normalize a Poisson distribution, also was conducted on each observation (Li 1964). Two additional transformations, the square root of each observation plus one half (Steel and Torrie 1980) and square root of each observation plus 1 (Snedecor 1956), were performed because of the large numbers of 0 and small counts for bare ground and individual species cover.

Sample size estimates were also obtained from our sampling data using the equation,

$$\hat{n} = \frac{t^2 pq}{(d)^2} \qquad (3)$$

where p = cover percentage, q = 100−p, and d = absolute error = 10%.

Results and Discussion

The sampling at Mandan was, by design, restricted to 4 cover components that are important for determining the adequacy of cover to protect the soil resource (Hofmann et al. 1983). The numbers of frames used to sample the 12 × 67-m plots represent

Table 1. Mean and variance of ground cover and estimated size of point frame samples required, based on 20, 40, 60, and 100 frames from mixed prairie.

Cover	Number of samples	Mean	Var	Equation 2[1]	Estimated size of samples Transformation $\sqrt{Y^2}$	$\sqrt{(Y + .5)}$	$\sqrt{(Y + 1)}$	Equation 3[3]
				----- Grazed -----				
Live	20	24.0	141	44	19	16	15	33
Live	40	28.5	208	44	16	15	14	35
Live	60	27.3	179	41	16	14	13	34
Live	100	26.2	173	42	16	14	13	33
Litter	20	71.5	150	6	2	2	2	36
Litter	40	68.8	257	10	3	3	3	37
Litter	60	70.3	210	8	2	2	2	36
Litter	100	72.0	190	7	2	2	2	34
Total	20	95.5	100	2	1	1	1	8
Total	40	97.2	56	2	1	1	1	5
Total	60	97.7	32	1	1	1	1	4
Total	100	98.2	31	1	1	1	1	3
Bare	20	4.5	100	869	627	196	131	8
Bare	40	2.8	56	1267	910	183	111	5
Bare	60	2.3	32	981	803	149	89	4
Bare	100	1.8	31	1594	1207	156	88	3
				----- Nongrazed -----				
Live	20	20.0	126	56	40	29	25	29
Live	40	16.5	152	95	55	39	33	24
Live	60	14.0	89	77	48	33	28	21
Live	100	17.7	167	89	51	37	32	25
Litter	20	78.5	203	6	2	2	2	30
Litter	40	83.0	160	4	2	2	2	24
Litter	60	85.8	89	3	1	1	1	21
Litter	100	80.7	180	5	2	2	2	26
Total	20	98.5	45	1	1	1	1	3
Total	40	99.5	5	1	1	1	1	1
Total	60	99.8	2	1	1	1	1	1
Total	100	98.4	24	1	1	1	1	3
Bare	20	1.5	45	3528	3528	228	122	3
Bare	40	.5	5	3313	3313	77	36	1
Bare	60	.2	2	10086	10086	33	14	1
Bare	100	1.6	24	1538	1293	147	81	3

[1]Estimated by $\hat{n}=t^2s^2/(\bar{x} * .10)^2$
[2]Y is the value of one observation.
[3]Estimated by $\hat{n}=t^2 pq/(10)^2$.

Table 2. Mean and variance of ground cover and estimated size of point frame samples required, based on 60 frames from grazed reclaimed and native pasture.

| | | | | Estimated size of samples | | | |
| | | | | | Transformation | | |
Cover	Mean	Var	Equation 2[1]	$\sqrt{Y}$[2]	$\sqrt{(Y + .5)}$	$\sqrt{(Y + 1)}$	Equation 3[3]
				Reclaimed			
Live	15.7	163	113	55	39	34	23
Litter	77.7	300	9	3	3	3	30
Total	93.4	226	5	2	2	2	11
Bare	6.6	226	855	562	215	153	11
Other	1.2	14	1713	1575	125	66	2
Brome	10.2	127	207	141	81	64	16
Alfalfa	3.0	32	589	529	133	83	5
Crested	0.5	8	5528	5109	88	42	1
Intermediate	0.8	11	2701	2462	113	56	2
				Native			
Live	29.2	187	38	12	11	11	35
Litter	67.7	249	10	3	3	3	37
Total	96.9	69	2	1	1	1	6
Bare	3.1	69	1165	772	180	113	6
Other	16.0	191	126	75	52	44	23
Needlegrass	4.8	66	476	369	133	92	8
Bluegrama	3.0	69	1286	951	197	121	5
Carex	5.3	49	290	251	102	72	9

[1]Estimated by $\hat{n}=t^2 s^2/(\bar{x} * .10)^2$.
[2]Y is the value of one observation.
[3]Estimated by $\hat{n}=t^2 pq/(10)^2$.

very concentrated sampling.

Industry and regulatory agencies often are interested in additional components of reclaimed mined land cover such as species composition. The samples from reclaimed mined land provided cover measurements typical of those taken to assess the adequacy of reclamation on mined land. These components usually have smaller means than live, litter, or total cover and equation (2) estimates a large sample size.

Using equation (2), the estimated number of samples required to adequately measure live, litter, total, and bare ground cover on small plots of mixed prairie, ranged from 1 frame (10 points) to over 10,000 frames (100,000 points) for various cover estimates using federal statistical requirements (Table 1). Total cover required fewest frames, only 1 to 2. Litter cover required 2 to 10 and live cover 41 to 95. Sampling to estimate percent bare soil required reading large numbers of frames with estimates ranging from 869 to 10,086 frames.

The number of frames required to adequately measure live, litter, total, and bare cover on reclaimed land ranged from 5 to 855 frames and from 2 to 1,165 frames on adjacent native pasture based on 60 samples (Table 2). The number of frames required to sample plant species range from 207 to over 5,500 frames (Table 2).

Table 1 illustrates the effect of frame numbers on the mean, variance, and the estimated sample size required to adequately sample live, litter, total, and bare ground cover. The means within a cover type were fairly consistent regardless of the number of frames used. However, not all means of cover comprising a small percentage of the total area were within 10% of each other. Increasing sample size, and increasing the number of samples did not improve the consistency of the means. This has been previously noted by Bonham (1982). The smaller the mean, the higher the calculated required sample size. Over 850 sample numbers were required to adequately measure the bare component of any of the sites which made up 0.2 to 6.6% of the ground cover (Tables 1 and 2). Estimates of individual plant species with small means also required large sample numbers.

An increase in the number of samples generally did not decrease variance nor was there any pattern of change in variance with

change in sample size (Table 1). Sowards (1983), using data analyzed from Trapper mine, also reported mean and variance did not change dramatically as sample size was increased beyond a certain number.

Bare soil is the complement of total cover because the sum of the 2 equals 100% ground cover for any particular area. Since both measure the same thing, the amount of cover or the lack of it, the variances of total cover and bare ground cover are identical. However, the same measurements used in equation (2) produce very different required sample sizes to measure the same cover. For example, the sample size, based on 60 frames, required to adequately measure total cover on the nongrazed mixture prairie area (Table 1) was only 1 frame whereas measurement of bare ground cover required up to 10,086 frames. The variance of each was 1.6667 but the means were very different, 99.8333 for total versus 0.1667 for bare. When equation (2) is solved, the numerator remains the same because t^2 and s^2 do not change, but the denominator changes greatly, which produces large differences in estimated numbers. This occurs for all total and bare cover measurements. Our data (Table 1) also illustrate that the mean and variance do not change dramatically as sample size is increased beyond a certain number, and unrealistic sample numbers are required when using equation (2) to estimate ñ. Results suggest that equation (2), as used, is inappropriate and impractical for use as a sample adequacy formula for reclamation assessment.

The adequacy formula is based on the assumption that the data are normally distributed. Histograms of the sampling distribution of live plant and litter cover data (Fig. 1A and 1B) tended towards the shape of normal distribution whereas bare ground (Fig. 1C) and individual plant species (Fig. 2A and 2B) were distributed towards a Poisson fashion (Li 1964). Histograms of total cover were mirror images of their bare cover counterpart. These latter data are very skewed and nonsymmetrical and would require some form of transformation or adjustment in order for the formula to be an efficient estimator of ñ (Bonham 1982).

Although the square root transformation is generally used to normalize the Poisson distribution, the mean must be larger than 10 to be effective, and then the transformation is not needed (Li
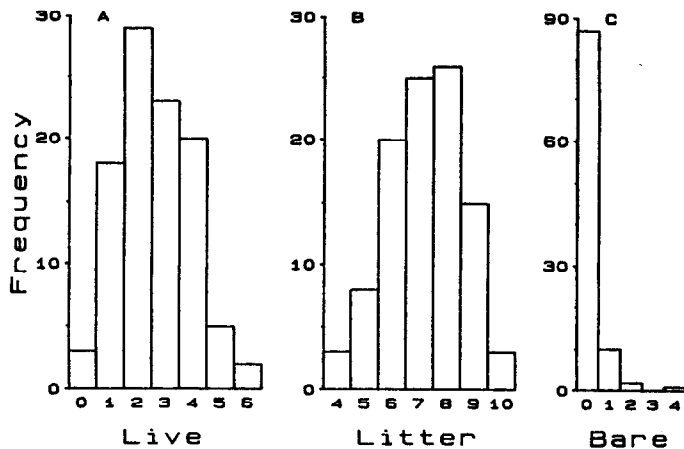
Fig. 1. Histogram of the distribution of live (A) and litter (B) cover and bare ground (C) sampled with 100 point frames on mixed prairie.
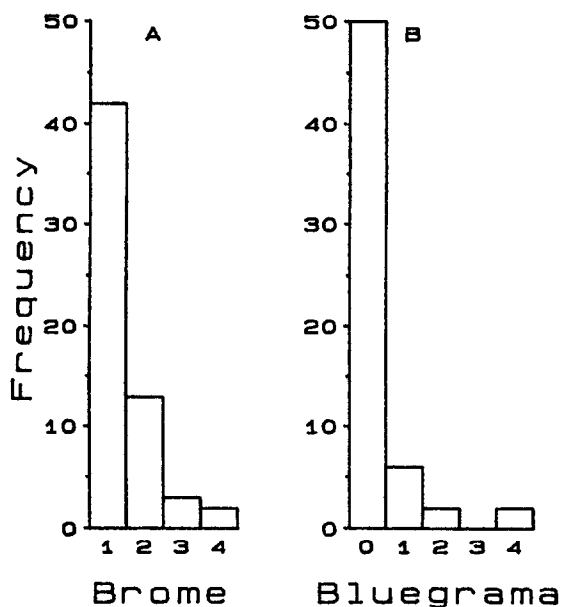

Fig. 2. Histogram of the distribution of smooth bromegrass (A) and blue grama (B) sampled with 60 point frames on reclaimed and native pasture.

1964). Our cover data in Table 1 suggested this is true. The square root transformation reduced the estimated number of frames required to sample live cover with means that ranged from the high teens to the mid 20's by about half. Estimated samples for bare means that were less than 5 were changed little and under 1.5 not at all.

Steel and Torrie (1980) recommended a $\sqrt{Y + .5}$ transformation (Y = the value of an observation) when some of the values are less than 15, especially when zeros are present, which describes our bare cover data. Snedecor (1967) suggested a $\sqrt{Y + 1}$ transformation if some counts are small. Both transformations have been applied to the data, Tables 1 and 2, which resulted in lower sample size estimates. Using Snedecor's transformation, numbers of frames required were reduced from unreasonable hundreds to less than 150. However, transformations such as these are generally preferred when data analysis is used to test hypotheses (Bonham 1989), not to determine sample adequacy.

Cochran (1977) lists 6 principal steps involved in the choice of sample size which are paraphrased below in the context of this study.

1. There must be some terms of desired limits of error. In cover sampling for reclamation, 90% has been defined.
2. Some equation that connects n with the desired precision of the sample must be found. Equation (2) has been suggested and is used in the preceding discussion to meet that need.
3. There will be parameters of the population that are unknown. These must be estimated. In equation (2) and similar equations, the variance usually is unknown. A preliminary sample is used to estimate the variance to compute the number of observations necessary. Additional observations are made and a new mean based on all the observations is computed.
4. Data often are to be published for major subdivisions of the population and desired limits of error are set up for each subdivision. Cover is composed of subdivisions such as live plant, litter, bare ground, rock, species, etc. For reclamation adequacy sampling, the desired limit of error has been set at 10% of the mean.
5. More than 1 item or characteristic is usually measured in a sample survey. If a degree of precision is prescribed for each item, the calculations lead to a series of conflicting values of n. Some method must be found to reconcile these values. When ($\bar{x}$ * .10) is prescribed as the limit of error, the degree of precision is changed depending on the abundance or absence of the cover component in the population. This causes a conflict because total cover and bare ground measure the same ground cover characteristic with the same variance, yet the number of samples required to measure either is very different.
6. The chosen value of n must be appraised to see whether it is consistent with the resources available to take the sample. Cover measurements requiring hundreds of point frames are costly, especially when improper equations are used to estimate the sample size. Since point data used to measure cover are proportion or percentage data and known to be a binomial or Poisson variate, a more appropriate equation to estimate sample size should be used.

When the proportion is equal or less than .30 or equal or greater than .70, the equation

$$n = \frac{t^2 PQ}{(kP)^2}$$

is adequate. P = cover percentage or proportion, Q = 100-P or 1-P, and kP = relative error. When the quantity to be estimated is a proportion of the population, PQ = S^2. Within the range of ≤.25, PQ is fairly stable. Outside this range, larger values of P will overstate the needed sample size (Cochran 1975).

When the population sizes are larger than 10,000, the size of the sample is important, not the percent of the population in the sample. The precision is then stated in terms of the population total (Cochran 1975). For equation (3) kP = d = absolute error = 10% maintains a constant degree of precision for each characteristic (Step 5 above) and meets the desired 10% of confidence width.

Sample sizes estimated using equation (3) ranged from 37 frames to 1 frame (Tables 1 and 2). Mean estimates determined by 40 samples were not improved by sampling with 60 or 100 frames (Table 1). Equation (3) also estimates the same sample size for total cover or its complement, bare ground.

Conclusion

This study provides a better understanding of the effects of sample numbers on point frame field data and applicability of the adequacy formula. It explains why Whitman and Siggeirsson (1954) suggested that many more points were needed to sample grassland vegetation than previous authors had suggested. Whitman and Siggeirsson used an adequacy formula whereas we and others used a number of different sample sizes to test the effect of

sample size on the mean.

Equation (2) is not applicable for estimating the number of point frames required to measure components of ground cover with means of less than 15% because they tended to be distributed in a Poisson rather than normal fashion. Equation (2) appeared adequate when cover means were greater than 15%. Cover components with means at 60% and greater require 10 or fewer frames. Since total cover and bare ground are complementary, equation (2) will estimate similar sample size only when they are at 50% ground cover. Using recommended data transformation did not correct non-symmetry although lower sample sizes were estimated.

Equation (3), which is appropriate for binomial or Poisson variable data, was also used to estimate sample sizes. Sampling more than the estimated size requirement did not greatly change mean values for various cover components, and satisfied Cochran's (1977) 6 principle steps for choosing sample size. It appears equation (3) is the most appropriate adequacy formula for estimating sample size required for ground cover measurements.

Literature Cited

Bonham, C.D. 1982. Approach to sample adequacy. p. 8-10. *In:* E.F. Aldon and W.R. Oaks (ed.). Reclamation of mined lands in the southwest. Proc. Symp. SCS, USDA Forest Serv., and WRCC-21, Albuquerque, New Mexico. 20–22 Oct. 1982.

Bonham, C.D. 1989. Measurements for terrestial vegetation. John Wiley & Sons. New York, N.Y.

Chambers, J.C., and R.W. Brown. 1983. Methods for vegetation sampling and analysis on revegetated mined lands. Gen. Tech. Rep. INT-151. Ogden, Utah. USDA Forest Serv.

Cochran, R.S. 1975. Sampling notes. Univ. Wyoming. Stat. Lab. #2014. p. 2-31–2-32.

Cochran, W.G. 1977. Sampling techniques. 3rd ed. John Wiley & Sons. New York, N.Y.

Crocker, R.L., and N.S. Tiver. 1948. Survey methods in grassland ecology. J. Brit. Grassl. Soc. 3:2-26.

Drew, W.B. 1944. Studies on the use of the point-quadrat method of botanical analyses of mixed pasture vegetation. J. Agr. Res. 69:289-297.

Hofmann, L., and R.E. Ries. 1989. Animal performance and plant production from continuously grazed cool-season reclaimed and native pastures. J. Range Manage. 42:248-251.

Hofmann, L., R.E. Ries, and J.E. Gilley. 1983. Relationship of runoff and soil loss to ground cover of native and reclaimed grazing land. Agron. J. 75:599-602.

Levy, E.B. 1933. Technique employed in grassland research in New Zealand. Imp. Bur. Plant Genet. Herb. Plants Bull. 11:6-16.

Levy, E.B., and E.A. Madden. 1933. The point method of pasture analyses. New Zealand. J. Agr. 46:267-279.

Li, Jerome, C.R. 1964. Statistical inference I. Edwards Brothers, Inc., Ann Arbor, Mich.

SAS Institute. 1985. SAS/GRAPH User's Guide. Version 5 Ed. Cary, N.C.

Snedecor, G.W. 1956. Statistical methods. Iowa State College Press. Ames, Iowa.

Steel, R.G.D., and J.H. Torrie. 1980. Principles and procedures of statistics. Second ed. McGraw-Hill Book Co., New York, N.Y.

Sowards, W.E. 1983. Achieving an adequate sample size for vegetation analyses. p. 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. Group, August 1983. CSU Range Science Series 35.

Whitman, W.C., and E.J. Siggeirsson. 1954. Comparison of line interception and point contact methods in the analysis of mixed grass range vegetation. Ecology 35:431-435.