Estimating Corn Lily Abundance in Bear Trap Meadow, July 2004

Mark Contois, Jameson Cahill, Nacha Chavez, Courtney Cacace, Gail Wechsler, Mariela Pauli, and Ellen Kosman

Faculty advisor: Prof. Ed Connor

Abstract

We analyzed the efficiency and accuracy of seven plotless density estimators gathered from Engeman *et al.* (1994) using field data on the density of Corn Lily (*Veratrum californicum*) in an alpine meadow of the Sierra Nevada Mountains. We found a significant correlation between the ranking of each estimator's variance and its efficiency as determined by the average time required to obtain a single sample estimate. This correlation shows that plotless estimators with lower variance require more time to apply in the field. For example, the AO3Q estimator had the smallest variance and took the most time to execute. We also used a plot-based sampling method (1 m^2 quadrats). The quadrat estimate showed a variance as low as that of the best plotless density estimator, and its efficiency was highest. We conclude that the adjusted efficiency of all seven plotless estimators is low in comparison to that of a plot-based method.

Introduction

Estimating the abundance of a plant species may seem at first glance a straightforward endeavor, but to obtain accurate estimates requires careful selection of an appropriate estimator and the allocation of an appropriate level of effort. We attempted to determine experimentally which, among a series of plotless density estimators and one plot-based estimator, would be best to obtain an accurate estimate of the abundance of a significantly aggregated meadow plant. To this end, we applied these estimators to a population of corn lily (*Veratrum californicum*) in Bear Trap Meadow ($39°38'.92$ N $/120°30'.98$ W), California, in July 2004.

Conducting numerous computer simulations, Engeman *et al.* (1994) used Monte Carlo methods to calculate the statistical properties of some 25 plotless and plot-based estimators for populations with different degrees of aggregation, including the angle-order quadrant (AO*n*Q), Kendall-Moran (KM*n*P), and ordered-distance (OD*n*C) families of estimators, as well as the variable-area transect (VAT). They concluded that the five best performing estimators in terms of minimized variance, when applied to aggregated populations, are the AO3Q, AO2Q, KM2P, OD3C and VAT (Table 1). We chose to examine four of these "best" estimators, excluding Kendall-Moran, which appeared to be the most complicated to execute in the field and required data other than distance measurements to the closest individual. In addition, since the measurements required by these estimators also suffice to estimate density under AO1Q, ODCI, and OD2C, we included these as well, giving a total of seven plotless estimators.

To determine whether estimators with higher variance could be superior to those with lower variance, we quantified the efficiency of each estimator by recording the average time required to obtain an estimate. We used the variance of each estimator as reported by Engeman *et al.* (1994), and obtained time measurements in the field. We also included the widely used quadrat sampling method as a baseline for comparison with the plotless estimators.

Table 1. Density estimators used in this study

Description*	Formula
Ordered distance (OD) estimators	
1. Closest individual (CI)	$\mathrm{ODCI} = (N-1)/(\pi\sum(R_{(1)i})^2)$
2. Second closest individual (2C)	$\mathrm{OD2C} = (2N-1)/(\pi\sum(R_{(2)i})^2)$
3. Third closest individual (3C)	$\mathrm{OD3C} = (3N-1)/(\pi\sum(R_{(3)i})^2)$
Angle-order (AO) estimators	
4. Closest individual per quadrant (1Q)	$\mathrm{AO1Q} = 12N/(\pi\sum(R_{(1)ij})^2)$
5. Second closest individual per quadrant (2Q)	$\mathrm{AO2Q} = 28N/(\pi\sum(R_{(2)ij})^2)$
6. Third closest individual per quadrant (3Q)	$\mathrm{AO3Q} = 44N/(\pi\sum(R_{(3)ij})^2)$
Variable area transect (VAT) estimator	
7. Variable area transect	$\mathrm{VAT} = (3N-1)/(w\sum\ell_i)$
Quadrat (QUAD) estimator	
8. Quadrat	$\mathrm{QUAD} = \sum q_i/(\ell_i w_i N)$

* AO: angle-order estimator; CI: closest individual; OD: ordered-distance estimator; QUAD: quadrat estimator; VAT: variable area transect estimator; 2C: second-closest individual; 3C: third-closest individual; ℓ_i = length searched from sample point to the gth individual (for VAT) or length of quadrat (for QUAD); N = sample size (number of sample points used in estimation); $R_{(g)i}$ = distance from the ith sample point to the gth CI; $R_{(g)ij}$ = distance from the ith sample point to the gth CI in the jth quadrant; w = width of strip transect; q_i = number of individuals counted in quadrat; w_i = width of quadrat

Reference: Engeman *et al.* (1994)

Methods

We used a satellite navigator (Magellan GPS Blazer 12) to create a map of Bear Trap Meadow (Figure 1) and selected sample points by generating random numbers with MATLAB Student Version 6.5; we first marked these points on the map of the meadow, and then located and marked them in the field using the navigator. We selected a total of 12 points, with 9 points sampled on the first day and 3 on the next day.

We formed two field crews, with three and four members, and assigned each sample point at random to a crew. Each crew was equipped with a quadrat (approximately 1.04 m^2, divided crosswise into four sections), a compass, a tape measure, a watch, and a 1 m length of twine.

Since *V. californicum* is highly clonal, we estimated the density of ramets in each case, and "individual" should be taken to mean "ramet" throughout.

At each sample point, we applied all eight estimators and recorded the time taken to obtain the necessary measurements for each estimator. We obtained quadrat estimates by centering the $\sim$1 m^2 quadrat on the sample point. We took angle-order measurements by using the four sections of the quadrat to represent the northeast, northwest, southwest, and southeast quadrants in which to measure distances to individuals. For both angle-order measurements and ordered-distance measurements, we measured the distance from the sample point to the first, second, or third nearest individual. We executed the variable area transect measurements by walking a radial line from the sample point at an angle randomly generated using MATLAB and using a 1 m length of twine, stretched perpendicular to the radial line, to identify the nearest individual falling within the 1 m wide transect.

We recorded the average time required to make the necessary measurements for each estimator at each sample point, assuming that the time required for computation was so small a fraction of the time spent in measurement as to be negligible.

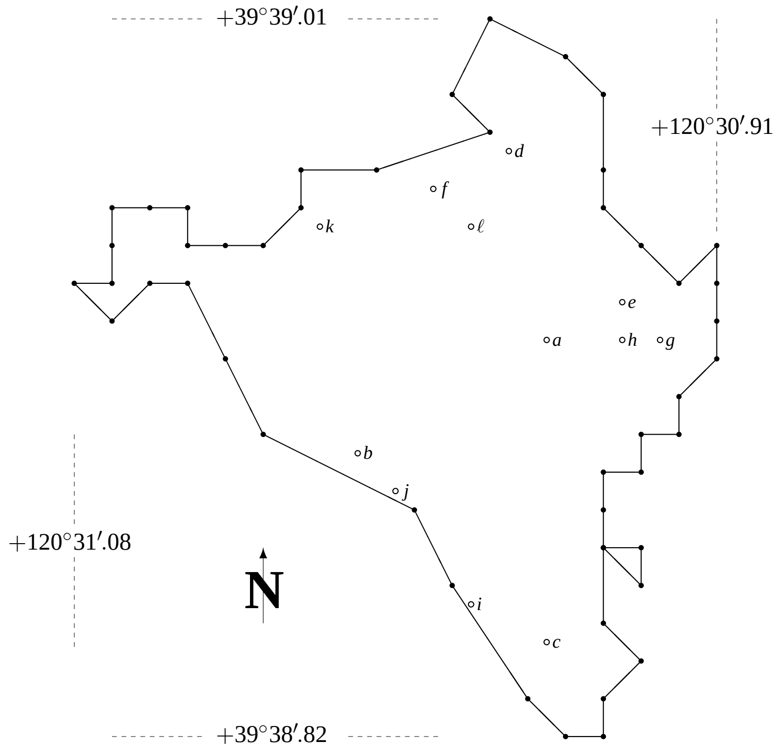
We used SPSS to assess the relationship between the average measurement time required for each estimator and the variance of the estimator as reported by Engeman *et al.* (1994), employing a single-tailed Spearman rank-order correlation test.

Analysis and discussion

Corn lily is a common plant in alpine meadows of the Sierra Nevada Mountains, tending to show highly aggregated patterns of distribution. In Bear Trap Meadow, corn lily is most common on the north side where it is found in extensive dense patches; it is also present throughout much of the meadow in scattered small clumps. Corn lily is scarce or absent in central areas of the meadow in the vicinity of a pond, and in drier, more upland parts of the meadow.

To determine which estimator was best in terms of accuracy adjusted for effort, we used the time required per sample point in combination with a measure of the variability of the estimator. We averaged the time spent per sample point among all sample points for each estimator (Table 2). Given that we did not know the true abundance of corn lily in the meadow, we relied on Engeman *et al.* (1994)'s estimates of variability of each estimator based on its ability to recover the known abundance of a simulated population

Figure 1. Bear Trap Meadow, Yuba Pass, California, as mapped using a satellite navigator. Solid circles represent marked boundary points; open circles labeled a through ℓ are sample points, randomly generated with a resolution of $0.01' \times 0.01'$. Note that the small triangular region connected at a single point was considered part of the meadow.



(Table 3), and chose to treat the quadrat data as our “best available” benchmark, based on the extremely low relative bias (RBIAS) estimates shown by the quadrat estimator in Engeman *et al*. (1994) even for aggregated populations.

Table 2. Average measurement time and rank ordering

Estimator	Time (seconds / sample point)	Rank
AO3Q	412.5	1
AO2Q	341.5	2
AO1Q	271.9	3
VAT	157.3	4
OD3C	75.1	5
OD2C	50.6	6
ODCI	34.5	7
Quadrat	*5.2*	–

Table 3. Rank ordering of estimators by Aggregate-15 relative root-mean-square error (RRMSE) with 10 sample points, from Engeman *et al.* (1994)

Estimator	RRMSE	Rank
AO3Q	0.24	1
AO2Q	0.33	2.5
AO1Q	0.59	6
VAT	0.33	2.5
OD3C	0.35	4
OD2C	0.45	5
ODCI	0.60	7
Quadrat	*0.26*	–

When the quadrat technique was omitted from the list of density estimators, we found that the ranking of estimators by time expenditure was significantly similar to the ranking by accuracy in the simulations by Engeman *et al.* (1994) (Spearman’s $\rho = 0.757$, $n = 10$, $\alpha = 0.049$; Figure 2). Including the quadrat sampling technique would have caused the Spearman test to fail, because the required time for quadrat sampling is the lowest of all methods while its theoretical accuracy is second only to that of AO3Q. Our conclusion, therefore, is that the accuracy of the plotless estimators we tested correlates strongly with their required time expenditure—but that the quadrat technique provides significantly better efficiency in terms of time spent—when applied to aggregated taxa such as *V. californicum*.

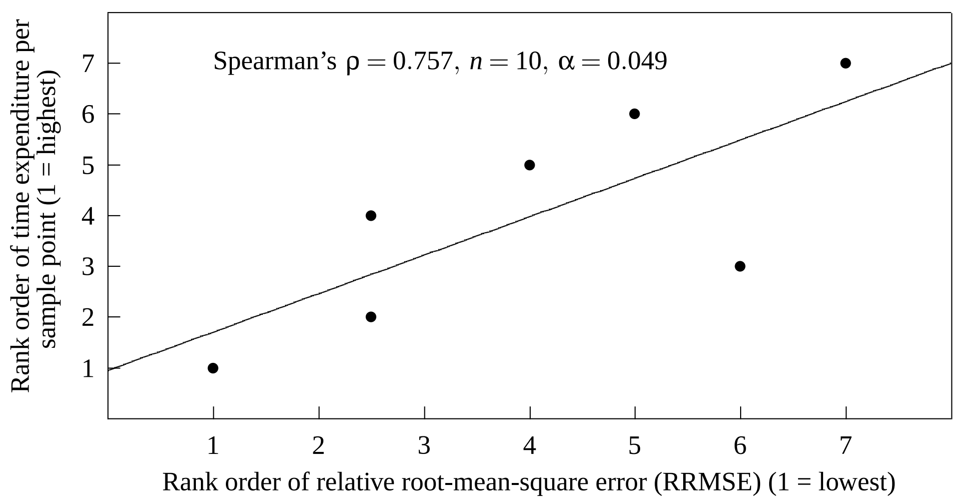
Although the purpose of this study was to test the efficiencies of selected estimators rather than to apply them in determining the density of corn lily in the meadow, we tabulated the experimental density estimates over all sample points nonetheless (Table 4), and it is striking to note the degree to which estimated densities varied among estimators—by nearly two orders of magnitude.

Table 4. Density estimates

Estimator	ramets / 10 000 m^2
ODCI ($n = 12$)	21.68
OD2C ($n = 12$)	44.74
OD3C ($n = 12$)	49.79
AO1Q ($n = 11$)	19.14
AO2Q ($n = 11$)	43.99
AO3Q ($n = 11$)	66.43
VAT ($n = 10$)	1288.
Quadrat ($n = 12$)	801.3

The statistical analyses in Engeman *et al.* (1994) indicate the tendency of plotless estimators to provide significant underestimates when applied to heavily aggregated populations, an observation that was borne out in our study with the single exception of the variable area transect. In Engeman *et al.* (1994), the VAT against Aggregate-15 ($n = 10$) gives mean RRMSE of 0.33 and RBIAS of -0.21; in our study the VAT estimate was approximately 161% of the quadrat estimate. One plausible explanation for this anomaly is psychological bias: if an individual is near the boundary of the

Figure 2. Relationship between the ranked variability of each estimator (from Engeman *et al.* (1994)) and the ranked average time spent per sample point.



transect, the experimenter may be more likely to count it “inside” than “outside,” a phenomenon that could easily lead to an overestimate.

Conclusions

We found that, among the plotless density estimators we tested, increased reliability in terms of lower variance correlated strongly with greater expenditure of effort in the field. The variance associated with the traditional technique of quadrat sampling was competitive with that of the most accurate plotless method, despite the quadrat’s requiring markedly less time to execute than any of the plotless estimators. Field biologists contemplating the use of plotless estimators for gauging the density of highly aggregated, sessile organisms such as *V. californicum* may find that greater efficiency is afforded by plot-based methods, including quadrat sampling.

References

[Engeman *et al.*] Engeman, R. M., Sugihara, R. T., Pank, L. F. & Dusenberry, W. E. 1994. A comparison of plotless density estimators using Monte Carlo simulation. *Ecology* **75**: 1769–1779.