

**ON USING ADAPTIVE CLUSTER SAMPLING DESIGN IN FOREST
INVENTORY: IT'S IMPLICATION TO FOREST BIODIVERSITY STATUS
REPORTING**

DENNIS P. PEQUE

College of Forestry and Natural Resources
Visayas State University
Visca, Baybay City, Leyte, Philippines

ABSTRACT

This paper presents adaptive cluster sampling (ACS) as a method of assessing forest biodiversity. In this study, ACS was used to estimate the abundance of ecologically sparse population of *Diospyros philippinensis* (Desrousseau) within the Visayas State University Forest Reserve. Its statistical efficiency using the Hansen-Hurwitz and Horvitz-Thompson estimators were analyzed by comparing them to the conventional systematic sampling (Syst) estimator.

Results indicated that adaptive cluster sampling (ACS) plots captured more trees into the sample compared to systematic sampling (Syst) plots. In addition, ACS estimates for mean and total numbers of individuals per ha was higher than systematic sampling estimates and in terms of variance ACS gave substantially lower variance than systematic sampling. However, the ratio of the adjusted SE of ACS to the adjusted SE of systematic sampling for each species and the combined data of the two species was generally lesser than 1 which means that ACS was not a better design than systematic sampling.

KEY WORDS: Adaptive cluster sampling, systematic sampling, Hansen-Hurwitz estimator, Horvitz-Thompson estimator, network.

INTRODUCTION

Forest biodiversity refers to the diversity within forest species, between species and of forest ecosystem. Forest biodiversity assessments provide information to support decision making options for the sustainable management and conservation of forest resources.

A number of sampling methods are available in assessing forest biodiversity but approaches could be different depending on the characteristics of the population studied and the variables or indicators assessed. They also differ in terms of their statistical performance.

Adaptive cluster sampling (ACS) is a relatively new sampling methodology which provides efficient unbiased estimators for the abundance of rare, clustered populations. The sampling procedure is data-driven such that when the information observed in the initial sampling units satisfy to some conditions to adapt, additional units are added until no more information are generated from the neighboring units or plots. It takes advantage of the clustering of individuals within the populations to adjust inclusion probabilities in favor of those plots harboring many individuals.

Since the introduction of ACS by Thompson in 1990 only few applications to forestry has been published so far (e.g. Roesch 1993; Acharya *et al* 2000; Talvitie *et al.* 2006 and Magnussen *et al.* 2005). While Roesch (1993) and Magnussen *et al.* (2005) evaluated the application and performance of ACS through simulation of empirical and artificial data, only the work of Acharya *et al.* (2000) and Talvitie *et al* (2006) used ACS literally in the field. Therefore, the current study is an addition to the very scanty applications of ACS to forestry.

METHODOLOGY

The study was conducted in the forest reservation of Visayas State University in Baybay City, Leyte, Philippines called Mt. Pangasugan. Specifically, the research site was conducted within Calbiga-a watershed covering an approximate area of 42.12 ha. It is located within the coordinates between $124^{\circ} 39' 48''$ to $124^{\circ} 39' 49''$ East and $10^{\circ} 45'$ to $10^{\circ} 46'$ North. The elevation ranges from 90 to 300 m a.s.l.

In this research, ACS and systematic sampling methods was used to estimate the abundance of sparse population of *Diospyros philippinensis* (Desr.) Gurke, a species belonging to the family Ebenaceae, and which has an *Endangered* (EN A1c) status (IUCN, 2000).

Systematic sampling (Syst) was used as the initial sampling design. After all plots were laid in the study site, adaptive cluster sampling procedure was then applied. Based on the size of the area and the size of the plots used, the distance between plots was 118.49 m.

Critical value C and the adaptive cluster sampling (ACS) process

Adaptive cluster sampling refers to designs in which the initial set of units is selected by some probability sampling procedure, and, whenever the variable of interest of a selected unit satisfies a given criterion; additional units in the neighborhood of that unit are added to the sample (Thompson 1992). Following this concept, a search for all trees were made to all neighborhoods of an initial systematic plot that met the minimum number of trees (of a certain species) and the process continued until no more trees were found. Here a cross pattern was used such that 4 plots are added one each from left, right, top and bottom to the initial plot that satisfied the condition to adapt (Figure 1).

The condition to adapt is that an initial plot should at least contain one individual tree ($C: > 0$) in any of the two species. All initial plots meeting this condition qualify for the application of adaptive scheme.

Thompson (1991) developed two ACS design utilizing both primary and secondary units. One of the designs involves initial simple random sampling without replacement (SRSWOR) of n_1 primary units where the units are formed from k systematic selections. These primary units contain an equal number of noncontiguous equidistant secondary units, all of which are sampled. Additional sampling is also done at the secondary unit level. This procedure was followed by Acharya et al (2000) in their studies to assess rare tree species in Nepal using ACS. In this study, however, systematic sampling was used as the design where $n_1=1$. Since an unbiased estimator of the variance is not available for this design (Thompson, 1991), the data was analyzed using the estimators of ACS with simple random sampling without replacement (SRSWOR) as the initial design.

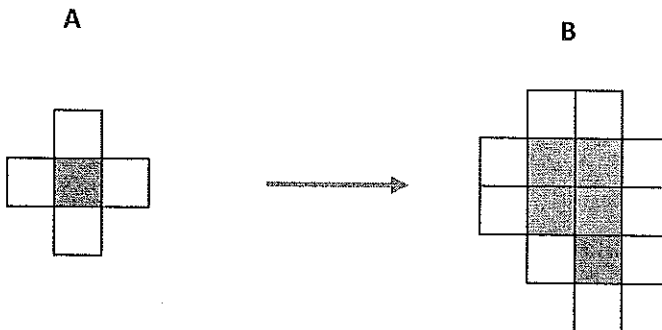


Figure 1. The adaptive sampling designs. *The gray-shaded unit indicates the initial seed plot that contained at least one species of interest. In figure A, the neighborhood is added top and bottom, left and right. If any of the added neighborhood likewise meet the condition to adapt, additional neighborhood are also made until they are surrounded by edge plots. In figure B, the bright-green-colored plots also meet the condition C to adapt.*

Three estimators of the population mean and variances are compared in this study (Table 1). The initial and the final samples (counts of all network quadrats plus the edge units) were computed using the SRS estimators whereas the ACS samples were computed using Hansen-Hurwitz and Horvitz-Thompson estimators.

In order to determine the efficiency of ACS estimators its variance estimates can be compared to the variance of systematic sampling. However, here the ration of standard of error of estimates from Syst and ACS was used as a method of evaluating the efficiency of ACS (Noon, et al., 2006). Accounting for the differences in the size of initial and final ACS samples, the variance of the initial sample was adjusted. Following Noon, et al., (2006), the formula is given below:

$$\hat{\sigma}_{SRS-adj}^2 = \left(1 - \frac{n'_{ACS}}{N}\right) \frac{\hat{\sigma}^2}{n'_{ACS}}$$

where n'_{ACS} is the number of units in the ACS sample (including edge units) and $\hat{\sigma}^2$ is the estimate of the population variance for systematic sampling.

RESULTS AND DISCUSSION

Physical attributes and number of trees encountered

Figure 2 shows the respective diameters (DBH) of eight *Diospyros philippinensis* trees encountered in the initial systematic plots and the 12 additional trees detected through ACS using a DBH limit of ≥ 7 cm. The overall average DBH and tree heights were 29.59 cm and 12.65 m, respectively.

Table 1. The estimators used in the study

Estimator	Mean	Variance
SRS (Initial sample)	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$	$\hat{S}_y^2 = \frac{1}{n-1} \left[\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i \right)^2}{n} \right]$
Hansen-Hurwitz	$\hat{\mu}_{HH} = \frac{1}{n} \sum_{i=1}^n w_i$	$\widehat{\text{var}}[\hat{\mu}_{HH}] = \frac{N-n_1}{Nn(n_1-1)} \sum_{i=1}^{n_1} (w_i - \hat{\mu}_{HH})^2$
Horvitz-Thompson	$\hat{\mu}_{HT} = \sum_{i=1}^K \frac{y_k^* I_k}{N\alpha_k} = \sum_{i=1}^K \frac{y_k^*}{N\alpha_k}$	$\widehat{\text{var}}[\hat{\mu}_{HT}] = \frac{1}{N^2} \left[\sum_{j=1}^K \sum_{k=1}^K \frac{y_j^* y_k^*}{\alpha_j \alpha_k} \left(\frac{\alpha_{jk}}{\alpha_j \alpha_k} - 1 \right) \right]$

It can be noted from the data that the encounter rate for *D. philippinensis* based on the number of plots assessed using systematic sampling and in ACS design was 25.81% and 30.76%, respectively. In the initial design, only five sub-plots contained at least one tree and when adaptive sampling was applied, six more adaptive plots (excluding edge and initial plots) were included which contains at least one individual. This suggests that this species clustered so much in this network thereby giving the ACS scheme the opportunity to increase the number of individuals that were detected.

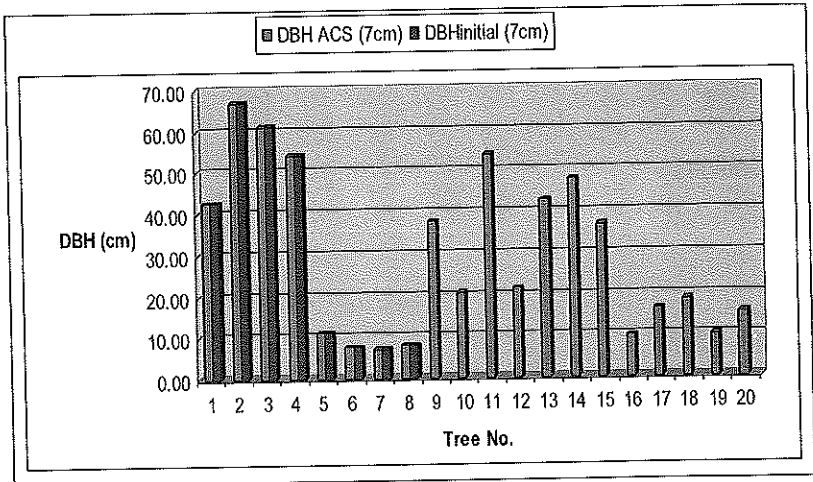


Figure 2. Diameters of *D. philippinensis* trees observed in the initial sample (8) final sample (12). Minimum DBH of trees considered was at least 7cm and the condition to adapt is at least presence of one tree.

Number and size of networks

Using a DBH limit ≥ 7 cm, there were five initial sampling units that detected the species satisfying the threshold value ($C: > 0$) and hence forming five networks of size at least one (depending on the number of objects detected and size of clusters that comprised a particular network). The total number of edge plots surrounding all the 5 positive networks was 28. The biggest cluster size comprising a particular network that exceeded the threshold value, $C: > 0$, was 14 and the smallest is five. A network is defined as a cluster without its edge units. Hence, the 3 initial plots that met the condition to adapt

yet surrounded by edge plots are considered a network each with size one. Therefore *D. philippinensis* samples has a total of 59 networks which is made up of 5 networks each with at least 1 unit and 54 networks of size 1 (28 edge plots and 26 initial plots). In terms of number of plots, the final sample contained 65 plots.

Estimates of *Diospyros philippinensis* and the efficiency of ACS estimators

The results in Table 2 indicate that the mean number of trees per ha using the four estimators did not vary too much. The two ACS estimators, Hansen-Hurwitz (HH) and Horvitz-Thompson (HT) estimators have roughly the same values with the estimated mean number of trees of the final sample computed using the simple random sampling (SRS) estimator. On the other hand, the mean number of trees per ha estimate of the initial sample, also calculated using SRS estimator, is slightly lower (around two trees) than the other estimators particularly ACS. This result suggests a gain in information when ACS is used instead of just the systematic sampling design. Conversely, the final sample (using ACS procedure) obtained the highest total number of trees among the estimators used.

According to Thompson (1992), the fact that the adaptive selection procedure produces high yield of observed objects gives the ordinary sample mean a tendency to overestimate. Thompson (1990) also noted that the sample mean \bar{y} corresponding to the final ACS sample are biased estimators of μ considering that each sampling unit has a different selection probability. Likewise, the sample variance s^2 corresponding to the final ACS sample is biased

for estimating σ^2 . The sample variance (as well as the total) of the initial and final samples were substantially higher compared to the variance estimates of ACS. As a result, this gives the ACS estimators a very low sampling error (e.g. SE % = 45.32%) compared to systematic sampling. Comparing HH and HT estimators, the latter gave the lowest variance estimates. This conforms to the findings, for example, of Thompson (1990), Christman (1997), Brown (2003), and Salehi (2003).

Table 2. Estimates of the number of *D. philippinensis* trees in the study area where C: > 0 and the minimum DBH of trees considered is 7 cm.

Estimates	ACS Producer (C: > 0)			
	Initial Sample	Final Sample	HH Estimator	HT Estimator
Mean (trees/ha)	11.47	13.68	13.19	13.32
Total No. of trees	483.05	576.00	555.51	560.87
Variance	917.38	1476.73	35.73	35.71
Total Variance	162751.34	2619858.76	63391.44	63347.41
Standard Error	1275.74	1618.60	251.78	251.69
Coefficient of Variation	264.10	281.01	45.32	44.87

In terms of statistical efficiency, ACS estimators were not efficient compared to systematic sampling by having a computed value of 0.617476 and 0.617686 for Hansen-Hurwitz and Horvitz-Thompson estimators respectively, which are less than 1.

IMPLICATIONS

Reporting status of biodiversity requires appropriate and efficient sampling method. The sampling design has to fit with the characteristics of the population under study to give a better estimate. Undoubtedly, it is evident that ACS captured more additional trees in adaptive plots which then improved its mean estimates compared to systematic sampling. This occurred in contiguous sampling units which were characterized by the presence of abundant objects (trees). Had it been that only systematic sampling was employed, then we can conclude that there were only 8 *D. philippinensis* trees within the study area instead of 20.

LITERATURE CITED

- ACHARYA, B., G. BHATTARAI, A. DE GIER, AND A. STEIN.** 2000. Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. *Forest Ecology and Management*. 137 (2000) 65-73.
- BROWN, J. A.** 2003. Designing an efficient adaptive cluster sample. *Environmental and Ecological Statistics* 10, 95-105.
- CHRISTMAN, M. C.** 1997. Efficiency of some sampling designs for spatially clustered populations. *Environmetrics*, 8, 145-66.
- GOLDBERG, N. A., J. N. HEINE and J. A. BROWN,** 2007. The application of adaptive cluster sampling for rare subtidal macroalgae. *Mar Biol* 151: 1343-1348.
- IUCN,** 2000. Red List- Philippines. www.iucnredlist.org/search-basic.html.

- LO, N. C. H., D. GRIFFITH, and J. R. HUNTER. 1997. Using a restricted adaptive cluster sampling to estimate Pacific hake larval abundance. *CalCOFI Rep.*, 38
- MAGNUSSEN, S., W. KURZ, and D. G. LECKIE. 2005. Adaptive cluster sampling for estimation of deforestation rate. *Eur. J. Forest Res.* 124: 207-220.
- NOON, B., N. M. ISHWAR, K. VASUDEVAN, 2006. Efficiency of adaptive cluster and random sampling in detecting terrestrial herpetofauna in a tropical rainforest. *Wildlife Society Bulletin* 34 (1): 59-68.
- PHILIPPI, T. 2005. Adaptive cluster sampling for estimation of abundances within local populations of low-abundance plants. *Ecology*, 86(5):1091-1100.
- ROESCH, F. A. 1993. Adaptive cluster sampling for forest inventories. *Forest Science.* 39(4): 655-669.
- SALEHI, M. M. 2003. Comparison between Hansen-Hurwitz and Horvitz-Thompson estimators for adaptive cluster sampling. *Environmental and Ecological Statistics* 10, 115-127, 2003.
- SMITH, D. R., R. F. VILLELLA, and D. P. LEMARIÉ, 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. *Environmental and Ecological Statistics* 10, 7-15.
- TALVITIE, M., O. LEINO and M. HOLOPAINEN, 2006. Inventory of sparse populations using adaptive cluster sampling. *Silva Fennica* 40(1): 101-108.
- THOMPSON, S. K. 1990. Adaptive cluster sampling. *Journal of the American Statistical Association*, 85, 1050-59.
- THOMPSON, S. K., 1991. Stratified adaptive cluster sampling. *Biometrika* (1991). 78 (2): 389-97.
- THOMPSON, S. K. 1992. *Sampling*. John Wiley and Sons.