Eur J Forest Res (2005) 124: 207–220
DOI 10.1007/s10342-005-0074-6

ORIGINAL PAPER

**Steen Magnussen · Werner Kurz · Don G. Leckie
Dennis Paradine**

Adaptive cluster sampling for estimation of deforestation rates

Received: 12 November 2004 / Accepted: 27 June 2005 / Published online: 10 August 2005

Abstract National estimates of deforestation rates may be based on a survey. Precise estimation requires an efficient design. When deforestation rates are low (< 1%) large sample sizes are required with traditional sampling designs to meet a precision target. This study explores the efficiency of adaptive cluster sampling (ACS) for this estimation problem. The efficiency is assessed by simulated ACS sampling from 18,200 × 200 km populations with 78–10,742 deforestation polygons (DFP) of different shape and size and average 10-year deforestation rates between 0.2% and 1.0%. Each population is composed of four million square 1 ha population units (PU) in a regular grid. Relative root mean square errors (RMSE) of ACS were, depending on sample size, 30–50% lower than comparable errors with simple random sampling (SRS) designs. ACS achieves this advantage by adaptively adding PUs to an initial SRS sample of size *n*. Realized ACS sample sizes were, on average, twice the nominal size (*n*). Three measures of ACS efficiency indicated that the costs of adaptively increasing the sample size are critical for the effectiveness of ACS. Population effects were manifest in all estimators. Estimates of the abundance, size, and shape of DFPs will allow a prediction of these effects. Populations dominated by a few large DFPs were clearly unsuited for ACS. The performance of ACS relative to that of SRS was similar across plot sizes of 1, 10, and 40 ha. The general conclusion of this study is that the lower RMSE of ACS remains attractive when the average cost of adaptively adding a PU to the initial sample is low relative to the average cost of sampling a PU at random.

Communicated by Michael Köhl

S. Magnussen (✉) · W. Kurz · D. G. Leckie · D. Paradine
Natural Resources Canada, Canadian Forest Service,
Pacific Forestry Center, 506 West Burnside Road,
Victoria, Canada V8Z 1M5
E-mail: steen.magnussen@nrcan.gc.ca
Tel.: +1-250-363-0712
Fax: +1-250-363-0775

Keywords Design efficiency · Network sampling · Expected sample size · Predicted efficiency

Introduction

National and regional estimates of totals or densities of rare species, events, or attributes in forested ecosystems are critically needed for a sustainable resource management, monitoring and control, and assessments of biodiversity, diseases, and ecological integrity. The permanent loss of forested lands to non-forest use caused by human activity is an example of a topical rare event (Koop and Tole 2001) with global significance (Fitzsimmons 2002; Stier and Siebert 2003). National estimates of deforestation areas are needed, primarily in countries that are signatories to the Kyoto Protocol (Dessai and Schipper 2003; Leaf et al. 2003). Local deforestation areas, henceforth called deforestation polygons (DFP) sum to national and regional deforestation areas. A periodic census of DFPs would deliver an estimate correct to within the bounds of measurement errors (Levy and Milne 2004), an ideal that is rarely possible; instead estimates can be obtained from some sort of sampling.

Choosing an effective sampling strategy for the estimation of the density of a rare species, event, or attribute is a challenge. Only reliable estimates provide the information needed for effective policies and management of forest ecosystem resources. Imprecise estimates can distort issues and policies if acted upon. Sample sizes needed to achieve a desired precision are often large and expensive (Green 1993; Madden and Hughes 1999; Christman 2000; Venette et al. 2002). Survey designers meet this challenge with stratified, sequential, or inverse sampling strategies (Christman 2000; Christman and Lan 2001; Su and Quinn 2003), or unequal probability sampling (Ståhl et al. 2000; Williams 2001a). The variance efficiency of these designs often depends critically on the population structure (Lo et al. 1997; Brown and Manly 1998; Acharya et al. 2000; Hanselman et al. 2003; Smith et al. 2003; Brown 2003).

Adaptive cluster sampling (ACS) has been suggested as an efficient sampling design for estimation of the population density of a rare attribute when the spatial distribution of the attribute is aggregated (Thompson 1990, 1991a, 1991b). An attribute has a spatially aggregated distribution when there is a positive correlation between the presence/absence of the attribute in spatially adjoining population units (PU) (Upton and Fingleton 1985). Here, the attribute is presence/absence of deforestation, the attribute value is the area of deforestation (DFA), and a PU is a square piece of land with a fixed constant area. Clearly, the spatial distribution of the attribute depends on the size of the PU and the size distribution of the DFPs. Each DFP is intersected by one or more spatially contiguous PUs, with the actual number depending on the PU area. The deforestation areas in each of the intersecting PUs are non-zero and their sum is equal to the area of the intersected DFP (within rounding). When the area of a PU is a fraction of the average area of a DFP, the spatial distribution of the attribute deforestation is clearly aggregated at the scale of the PU.

The concentration of DFPs in and around centers of human economic activity may create another aggregated spatial distribution of DFPs at the scale of economic activity zones (Sader et al. 2001; Srivastava et al. 2002; Fitzsimmons 2002; Evelyn and Camirand 2003; Fuller et al. 2004). The purported efficiency of ACS is for populations where spatial aggregation of the attribute is at the level of PUs. In this study, we assess the efficiency of ACS when the size of a PU is about one half the size of an average DFP.

In ACS, the data collection procedure depends on the data collected. Data collection begins with a simple random sampling (SRS) of a fixed number of PUs. In locations where the attribute is observed in a sampled PU, additional PUs are added to the sample according to a specific set of rules. The benefit is a larger yield of sampled PUs with the rare attribute and consequently more precise estimates of population parameters associated with the attribute—a benefit that comes at the expense of an uncontrolled increase in the number of sampled PUs. The magnitude of the increase will depend on the spatial distribution of the attribute (Lo et al. 1997; Acharya et al. 2000; Hanselman et al. 2003; Smith et al. 2003; Brown 2003). Capping the number of PUs sampled by a stopping rule has proven difficult and invariably introduces a bias in estimates of population parameters (Mohammad and Seber 1997; Morgan 1997; Brown and Manly 1998; Christman and Lan 2001; Pollard et al. 2002; Su and Quinn 2003).

Practical experience with ACS for estimating the density of a rare forest resource attribute remains scant. To our knowledge only two studies have been published: one for the estimation of the density of three rare tree species in a 40-ha forest in Nepal (Acharya et al. 2000), and the other for a simulated point sampling of tree species densities with a variable plot radius within a 3.1 ha mixed species stand in Maine (Roesch 1993). Both confirmed the potential (variance) efficiency of ACS but also pointed to the risk of cost overruns due to the adaptive increase in sample size. To assess the effectiveness of ACS, relative to that of SRS, for a regional (national) estimation of deforestation rates, this study compares bias, root mean square errors (RMSE), and variance efficiency (E) of ACS and SRS in simulated sampling. Sampling is from 18 artificial populations representing an important range of presumed variation in sizes, shapes, and spatial distribution of DFPs in Canada. We also explore the potential for predicting the sampling variance and thus the efficiency of ACS from descriptive statistics of DFP attributes.

Materials and methods

Populations

Eighteen artificial spatial populations of DFPs were studied. Each population covers an area of $200 \times 200\ \text{km}^2$. The number, area, shape, and spatial distribution of DFPs in these 18 populations are presumed to mirror an important range of concurrent regional decadal deforestation patterns in Canada. The patterns were derived from a variety of sources, including actual decadal deforestation events obtained from multi-temporal analysis of satellite images and aerial photography, known landscape patterns, and expert knowledge. Each population contains between 78 and 10,742 DFPs (median 1,540). The outline (shape) of a DFP is "rectangular," "linear," or "irregular."

DFPs were placed at random inside the population area with a correction for overlap. A fraction (0.2, 0.4, 0.5, 0.6, 0.7, 0.9, or 1.0) of the DFPs were placed inside square areas with a higher than average density of DFPs. The size of the high-density areas varied from 25 to 2,500 km^2 and in numbers from 2 to 243 per population with a marked tendency to an inverse relationship between their frequency and size. Table 1 summarizes statistics on the number and areas of DFPs. A sample of representative DFPs is in Fig. 1. The DFPs in this figure have been scaled for display purposes only. The spatial and area distribution of DFPs in four randomly selected populations are portrayed in Fig. 2.

Defining population units

A population is assumed to exist in the form of a classified binary (deforestation, no deforestation) image. In practical applications, only the outline of the population and the sampled parts of the image will be known; non-sampled parts remain, in principle, unknown. The sampling design for estimation of population-specific deforestation rates can be based on a point sampling of

Table 1 Population (POP) rates of deforestation (μ_{DF}), number of deforestation polygons (DFP), median (DFA_{med}), mean (DFA-mean), and maximum (DFAmax) polygon deforestation areas (unit: ha). The last two columns are the average number of population units per network of deforestation units ($\bar{n}_{\Psi_{DFA}}$) and the average perimeter-to-area ratio $\bar{R}_{P:A}$ of the DFP, respectively. See text for details

POP	μ_{DFA}%	DFP	DFA_{med}	DFA_{mean}	DFA_{max}	$\bar{n}_{\Psi_{DFA}}$	$\bar{R}_{P:A}$
1	0.25	2,604	2.2	3.7	87.7	21	5.4
2	0.49	5,258	2.2	3.6	92.8	22	5.8
3	0.96	10,742	2.1	3.4	93.7	24	6.3
4	0.24	1,430	1.8	6.1	186.0	25	3.9
5	0.49	2,857	1.8	6.2	249.2	27	5.4
6	0.98	5,796	1.8	6.2	202.7	27	4.6
7	0.23	2,348	1.9	3.9	108.9	34	17.9
8	0.45	4,774	1.9	3.8	121.1	35	17.3
9	0.19	1,911	2.0	4.5	82.9	59	39.9
10	0.24	768	6.7	11.5	147.4	67	28.8
11	0.49	1,524	7.0	12.0	203.7	69	24.7
12	0.94	3,143	6.2	11.1	222.1	67	21.8
13	0.22	78	0.7	102.9	6,586.4	227	3.2
14	0.44	162	0.7	100.8	9,002.7	259	14.3
15	0.88	341	0.7	94.9	7,606.2	262	49.9
16	0.24	1,540	2.7	5.8	185.7	24	2.2
17	0.49	310.4	2.7	5.7	242.6	24	2.8
18	0.98	6,259	2.5	5.7	137.3	25	2.5
Mean	0.48	3,036	2.6	54.7	1,420.0	74	14.3

image locations (each population is composed of an infinite number of points), or on a sampling frame which uniquely subdivides the image of the population into a finite set of units of known size, shape, and location (Särndal et al. 1992). We chose the latter as ACS is conveniently suited for implementation when the sample frame is fixed. For the sample frame to be fixed, the unit size, the unit shape, and the unit locations must be uniquely defined for all units in the population. Practical and statistical considerations and context dictates the choice of units. A frame constructed by a tessellation of the image of a population into a regular grid of square cells appeals on grounds of simplicity, convenience, and congruence with remotely sensed imagery.

In the current study, there is no natural subdivision of the population into PUs. Given that ACS can only be efficient if a majority of the DFPs have an area larger than that of a PU (Thompson 1990), we chose a unit size that would be smaller than the median DFP of all 18 populations (2.1 ha) but also larger than the average of the smallest 5% DFPs (0.51 ha). We chose—mainly for reasons of computational convenience—a unit size of 100×100 m^2 (1 ha). Although a survey designer would not have access to an area distribution of DFPs, subject knowledge would exist to guide the choice of PU. A preliminary study indicated that the relative performance of ACS and SRS, in terms of variance efficiency, was only slightly affected by the choice of PU size, at least as long as the area of a square PU was between 0.25 and 4 ha. With the choice of a 1-ha square PU, the sample frame becomes a regular grid of $N = 4 \times 10^6$ PUs in 2,000 rows and 2,000 columns.

The DFA in each of the 4 million PUs was determined by intersecting the DFPs with the PUs. Partitioning of the area of a DFP to PUs was done to an accuracy of ± 5 m^2. After partitioning the total deforestation areas to individual PUs, a calibration of PU deforestation areas was done to ensure that their total matched that of the total of the DFP areas. At the scale of a PU, the size of a DFP is defined as the number of PUs intersecting the DFP. The deforestation rate of a PU is the 10-year deforestation area (unit: ha) in the PU divided by the area of the PU (1 ha).

Sample plots

Sample plots are composed of one to several spatially contiguous PUs. We assess the performance of plots composed of m PUs (m = 1, 10, or 40). The multi-unit plots are columnar plots formed by stacks of 10 and 40 PUs. They emulate adaptive line sampling. It is known that for a given plot area, a long narrow plot is more (variance) efficient than square plot of the same size (Correll and Cellier 1987; Magnussen 2001). For each plot size (m), the number of plots M in the population is $M = N \times m^{-1}$ or, specifically, 4×10^6, 25×10^4, and 10^5.

Sampling objectives and sampling attribute

Our sampling objective is the estimation of population-specific 10-year rates of deforestation (μ_{DFA}) defined here as the total 10-year deforestation area divided by the area of the population. In our study, μ_{DFA} is equivalent to a density estimate since the DFPs embody the 10-year total of deforestation. In the context of this study, the terms density and rate are exchangeable.

Let y_{ij} denote the DFA (unit: ha) in the jth PU in the ith plot (j = 1, ..., m; i = 1, ..., M). An unbiased estimator of μ_{DFA} is the expected value of y_{ij} over all PUs divided by the area of a PU (|PU| ≡ 1 ha).

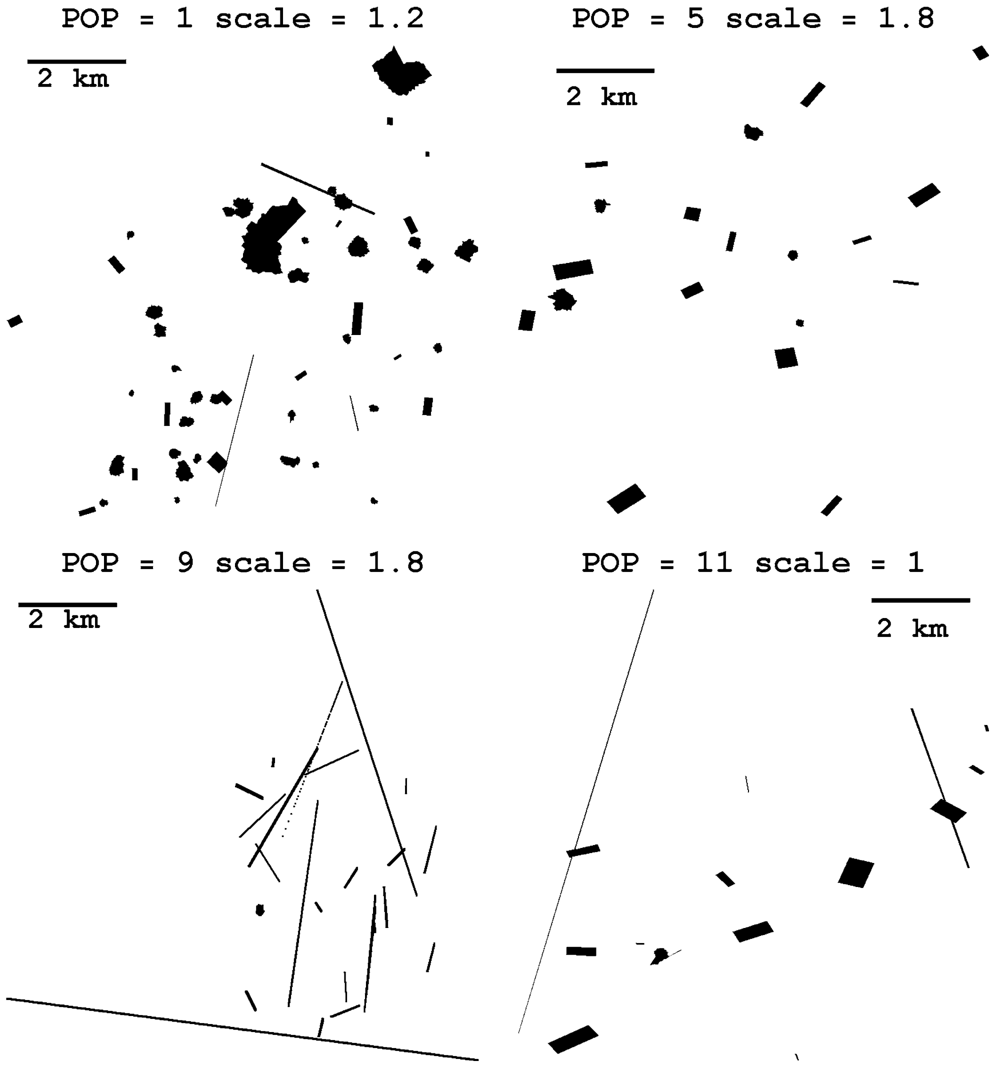

Fig. 1 Examples of deforestation polygons (DFP) in four randomly selected populations (POP). The DFPs have been scaled in size while centered at their actual location to allow all DFPs in the area to be displayed. Shown is the 10 km × 10 km quadrant with the maximum number of DFPs

Simple random sampling designs and estimators

SRS without replacement with a sample (s) of n plots of size m PUs was simulated in each population. Sample sizes (n) were 200, 400, ..., 3,000 for $m = 1$; 100, 150, ..., 800 for $m = 10$; and 50, 75, ..., 400 for $m = 40$. These designs reflect an anticipated trade-off between the number of plots and number of units sampled. As plot size increase fewer plots are sampled. Sampling within each population, plot size, and sample size was repeated 400 times. An unbiased design-consistent SRS estimator of the true μ_{DFA} is (Cochran 1977)

$$\hat{\mu}_{\mathrm{DFA}} = \frac{1}{n} \times \sum_{i \in s(n)} \bar{\bar{y}}_i, \tag{1}$$

where $\bar{\bar{y}}_i$ is the mean DFA per unit area in the ith plot. An unbiased estimator of the sampling variance of $\hat{\mu}_{\mathrm{DFA}}$ is (Cochran 1977)

$$\widehat{\mathrm{var}}\left(\hat{\mu}_{\mathrm{DFA}}\right) = \frac{(1 - n/M)}{n(n-1)} \times \sum_{i \in s(n)} \left(\bar{\bar{y}}_i - \hat{\mu}_{\mathrm{DFA}}\right)^2. \tag{2}$$

The mean of $\hat{\mu}_{\mathrm{DFA}}$ over the 400 repeat samples is denoted $\hat{\bar{\hat{\mu}}}_{\mathrm{DFA}}$, and the replicate variance of $\hat{\mu}_{\mathrm{DFA}}$ is $\mathrm{Var}(\hat{\mu}_{\mathrm{DFA}})$. From theory, we know that the expected value of the SRS estimates $\hat{\mu}_{\mathrm{DFA}}$ will equal the true population value; however, after just 400 repeat samples we can only expect that $\hat{\bar{\hat{\mu}}}_{\mathrm{DFA}} \cong \mu_{\mathrm{DFA}}$. Any discrepancy between $\hat{\bar{\hat{\mu}}}_{\mathrm{DFA}}$ and μ_{DFA} is due to sampling error, which should be accounted for when comparing an SRS estimator with an ACS estimator. The RMSE of $\hat{\bar{\hat{\mu}}}_{\mathrm{DFA}}$ in Eq. 3 is therefore used for design comparisons (Cochran 1977)

$$\mathrm{RMSE}_{\mathrm{DFA}} = \hat{\hat{\delta}}^2_{\mathrm{DFA}} + \mathrm{Var}\left(\hat{\mu}_{\mathrm{DFA}}\right), \tag{3}$$

where $\hat{\hat{\delta}}_{\mathrm{DFA}} = \hat{\bar{\hat{\mu}}}_{\mathrm{DFA}} - \mu_{\mathrm{DFA}}$. In the comparison between SRS and ACS, we use the term bias for $\hat{\delta}_{\mathrm{DFA}}$.

Adaptive cluster sampling

ACS begins with a SRS of n plots of size m from the M plots in the population (Thompson 1990, 1991a, 1991b), where n and m take the same values as outlined for SRS.

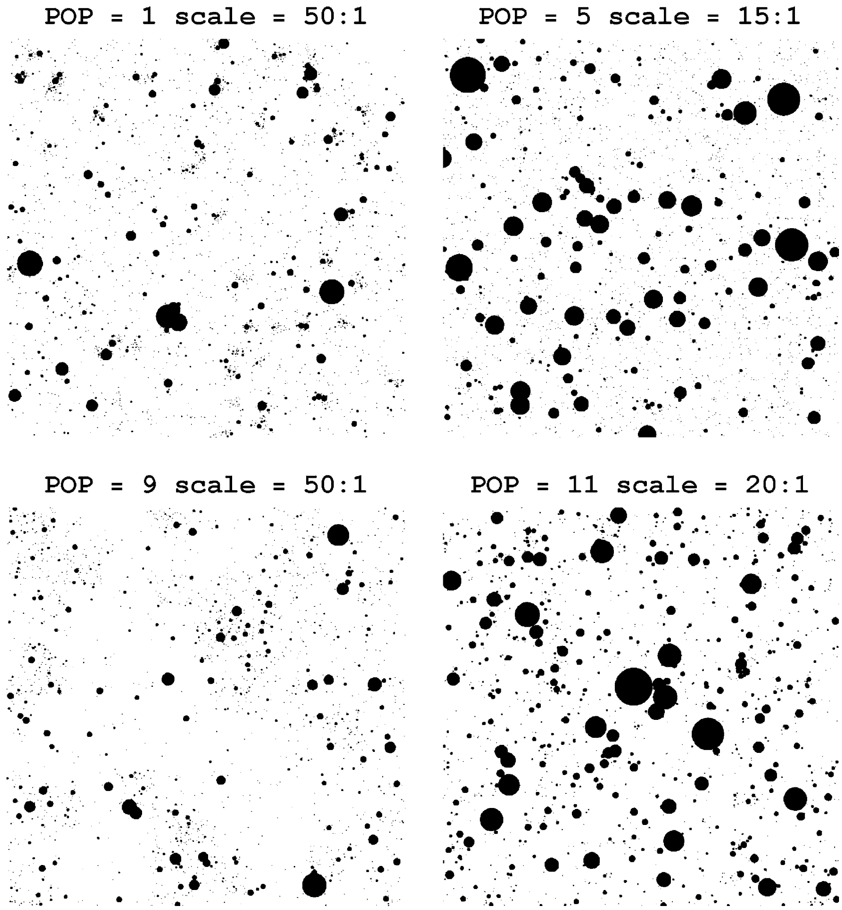

Fig. 2 Generalized deforestation polygons (DFP) in four randomly selected populations (POP). Each DFP is drawn as a circle scaled to an area suitable for the display of location and relative DFP sizes. Scaling factors are given in each subplot

Each sample plot is then adaptively expanded in size according to a rule that applies whenever certain conditions are met. The rule of expansion can be varied freely; we opted for the rule given in Thompson (1990) which is quite easy to implement although it may at first appear cumbersome. Each PU in the SRS sample is queried about whether it contains a positive deforestation area ($y_{ij} > 0$) or not. In the affirmative case, the four adjacent PUs (above, below, left, and right) are added to the plot in which the PU with the positive DFA resides. A union of the originally sampled PUs and the adaptively added PUs then creates n plots with size $\geq m$. Hence, no PU is added more than once to a specific sample plot. This process is iterated until the plot sizes remain constant between iterations. Plot sizes will cease to increase when the deforestation area in all PUs located next to the outside edge of the expanded plots is zero.

From this adaptive rule, it follows that whenever a PU in the initial SRS sample intersects with a DFP that all PUs intersecting this DFP will be included in the sample. Also, any DFP separated by less than one PU from a DFP intersected by the initial SRS sample is included as well. Two DFPs separated by less than one PU are connected. Thus, a single DFP may be included in more than one expanded sample plot. In our populations this rarely happened.

The number, size, shape, and spatial distribution of DFPs are clearly important factors determining the adaptive expansion of sample plots. A graphical illustration of the adaptive plot expansion is given in Fig. 3 for plot sizes of one and ten PUs. A DFP with a DFA of 3.57 ha is portrayed as a black rectangle. Nine PUs intersect the DFP, and they are displayed in a dark-gray tone. Any initial sample with plots of size one that includes one of the nine dark-gray PUs will be adaptively expanded to include the 21 PUs outlined by a dashed line (Fig. 3 top). An example of the adaptive expansion of a 1×10 PU plot when it intersects the same DFP is in the lower half of Fig. 3 where the plot is expanded to 27 PUs.

Networks of PUs with deforestation

To appreciate the following presentation of ACS estimators of μ_{DFA} and sampling variance, it may be helpful to regard a population as composed of three types of PUs. One type is the PUs with $y_{ij} = 0$ and not adjacent to a PU with $y_{ij} > 0$. This type of a PU is only sampled if they are part of the initial SRS sample; they do not trigger any adaptive expansion of the sample. The probability of including this type of PU in the original sample is obtained by first principles (Cochran 1977).

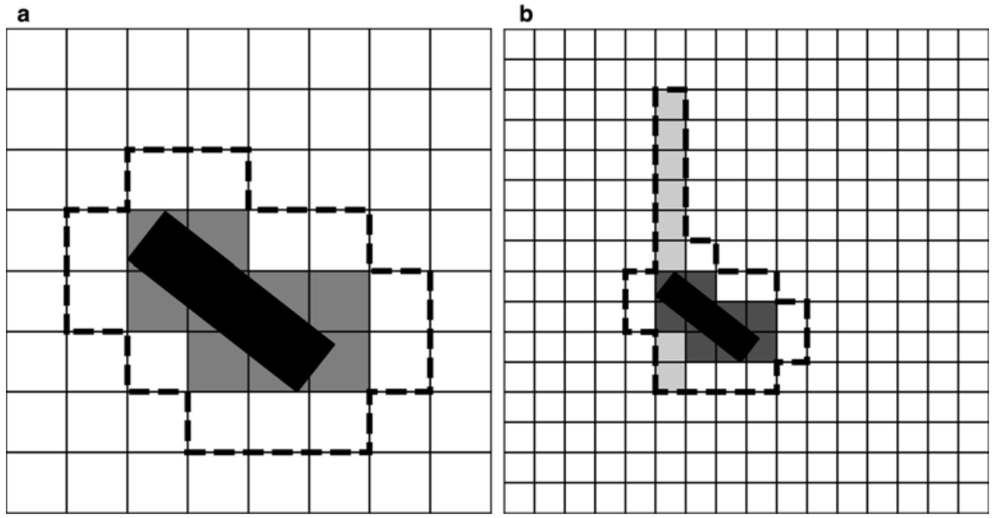

Fig. 3 Examples of adaptive plot expansions. **a** $m = 1$. The *black rectangle* is a deforestation polygon (population# 1, polygon# 2,153) located in a regular array of 1 ha population units. The associated deforestation network of units with a positive deforestation area is in a dark-gray tone. The outline of an adaptively expanded plot is indicated by *dashed line*. **b** $m = 10$. Initial SRS sample plot is in light-gray tone. See text for details

The second type has a DFA of zero but is adjacent (edge) to a DFP. This type is selected if intersected by the initial sample or if it is adjacent to a DFP intersected by one or more PUs in the initial sample. Selection probabilities of this second type depend on the size of adjacent and connected DFPs. The third type are the PUs intersecting one or more DFPs. Given the adaptive protocol, all PUs intersecting a DFP or a set of connected DFPs have the same inclusion probability, which is proportional to the size of the intersected DFP plus all DFPs connected to it.

The PUs intersecting a DFP or a set of connected DFPs are collectively referred to as a deforestation network (Psi). The size of a deforestation network is the number of PUs in the network. Selection of one PU in a deforestation network leads to the selection of all other PUs in the network. All PUs in a deforestation network are connected. A PU with $y_{ij} = 0$ is a non-deforestation network of size one. A population can thus be viewed as consisting of, say, K distinct non-overlapping networks.

ACS can be viewed as a SRS of networks. Each sample plot will intersect with one or more networks, a majority of which will be non-deforestation networks of size one. The number and size of sample plots determines the probability of intersecting a network. Since all PUs in a network intercepted by an initial SRS plot are included in the final ACS sample, it follows that the DFA per PU value for a specific plot type and location is that of the associated network(s). Said otherwise, in ACS the within-network variance of DFA per PU has been eliminated from the sample variance. ACS sampling variance is therefore exclusively due to the among-network variance in DFA per PU. From this follows that the (variance) efficiency of ACS will depend on the relative magnitude of the within-network variance of DFA per PU. Networks of size one have, by definition, a within-network variance of zero.

ACS estimators of deforestation rates

We present two ACS estimators of the deforestation rate (density) and sampling variance, the modified Hurwitz–Hansen (HH) estimators and the modified Horvitz–Thompson estimators (HT). The modified HH estimators view ACS as SRS of network deforestation rates (densities). The modified HT estimators exploit the post-sampling estimates of the inclusion probabilities of the networks composed of the three types of PUs. These inclusion probabilities, in turn, are used for the derivation of "classical" HT estimators for sampling with unequal inclusion probabilities (Thompson 1990). Both the HH and the HT estimators are asymptocally design-unbiased, that is, as sample size increase the expectation of a sample estimator approaches the true population value. Although the modified HH estimators can be less efficient than the modified HT estimators (Lo et al. 1997; Christman 2002) they are said to be less sensitive to the spatial structure of the population attribute (Salehi 2003).

The modified HH estimator of μ_{DFA} is

$$\hat{\hat{\mu}}_{\text{DFA.HH}} = \frac{1}{n} \times \sum_{i=1}^{n} \sum_{k=1}^{K} \delta_{ik} \times \mu_{\text{DFA.}\Psi_k}, \tag{4}$$

where $\mu_{\text{DFA}}.\Psi_k$ is the deforestation rate (density) in the kth network, and δ_{ik} is an indicator value taking the value of 1 if the kth network is intersected by the ith sample plot and zero otherwise. A plot of size one ($m = 1$) can only intersect one network, but a plot of size m ($m > 1$) can intercept more than one network. All plots intersecting a specific network contribute the same deforestation rate (density) value to the modified HH estimator in Eq. 4. Equation 4 applies to any plot shape and size that completely tessellates the population. The effect of plot size is through δ_{ik}. For a fixed sample size n, increasing m simply increases the number of intercepted networks. Note, the HH estimator in Eq. 4 appears to be an un-weighted mean of rates. Recall, however, that networks are sampled with probability proportional to their size (number of PU) and that the total deforestation area in a network is size multiplied by $\mu_{\mathrm{DFA}}.\Psi_k$. Hence, network size cancels out of the estimator. The associated estimate of the sampling variance of $\hat{\hat{\mu}}_{\mathrm{DFA.HH}}$ is

$$\widehat{\mathrm{var}}\left(\hat{\hat{\mu}}_{\mathrm{DF.HH}}\right) = \frac{(1-n/M)}{n(n-1)} \times \sum_{i=1}^{n}\sum_{k=1}^{K} \delta_{ik} \times \left(\hat{\hat{\mu}}_{\mathrm{DFA}.\Psi_k} - \hat{\hat{\mu}}_{\mathrm{DFA.HH}}\right)^2. \tag{5}$$

We also computed the average of HH estimates of deforestation in repeat sampling $\bar{\hat{\hat{\mu}}}_{\mathrm{DFA.HH}}$, bias $\hat{\hat{\delta}}_{\mathrm{DFA.HH}}$, and $\mathrm{RMSE}_{DFA.HH}$ as outlined for SRS and the extension of notation. The modified HH estimators in Eqs. 4 and 5 are essentially well-known SRS estimators if one considers the network deforestation rates (density) as the attribute of interest.

The modified HT estimators are derived from the probability of having an observation from the kth network in the sample; accordingly the modified HT estimator of the deforestation rate (density) is

$$\hat{\hat{\mu}}_{\mathrm{DFA.HT}} = \frac{1}{m \times M} \times \sum_{k=1}^{K} \frac{y_{\Psi_k} \times \delta_k}{\alpha_k}, \tag{6}$$

where y_{ψ_k} is the total deforestation area in the kth network, α_k is the probability that network k will be in the sample, and δ_k is an indicator value taking the value 1 if the kth network is intercepted by the initial random sample and zero otherwise. The probability of having the kth network intercepted by the sample is

$$\alpha_k = 1 - \binom{M - x_k}{n} \times \binom{M}{n}^{-1}, \tag{7}$$

where x_k is the number of plots (of size m PUs) that intersect network k. An unbiased modified HT estimator of the variance of $\hat{\hat{\mu}}_{\mathrm{DFA.HT}}$ is

$$\widehat{\mathrm{var}}\left(\hat{\hat{\mu}}_{\mathrm{DFA.HT}}\right) = \frac{1}{M^2} \times \sum_{k=1}^{K}\sum_{h=1}^{K} \frac{y_{\Psi_k} \times y_{\Psi_h} \times \delta_k \times \delta_h(\alpha_{kh} - \alpha_k\alpha_h)}{\alpha_k\alpha_h\alpha_{kh}}, \tag{8}$$

where α_{kh} is the probability that the initial sample intersects both network k and h at least once.

$$\alpha_{kh} = 1 - \left\{\binom{M-x_k}{n} + \binom{M-x_h}{n} - \binom{M-x_k-x_h}{n}\right\} \times \binom{M}{n}^{-1}, \tag{9}$$

with the convention that $\alpha_{kh} = \alpha_k \times \alpha_h$ if $h = k$. Again, we see that the effect of plot size is on the network inclusion probabilities (M declines with increasing plot size).

In ACS, both $\hat{\hat{\mu}}_{\mathrm{DFA.HH}}$ and $\hat{\hat{\mu}}_{\mathrm{DFA.HT}}$ rely on "realized inclusion probabilities" since the true inclusion probability of a unit (plot) is not known in advance of sampling. As such the estimators are asymptotically unbiased.

Estimators $\bar{\hat{\hat{\mu}}}_{\mathrm{DFA.HT}}$, bias $\hat{\hat{\delta}}_{\mathrm{DFA.HT}}$, and $\mathrm{RMSE}_{\mathrm{DFA.HT}}$ were obtained as outlined above for SRS.

Among- and within-network variance of deforestation rates

In ACS, the attribute recorded for each of the initial n SRS plots is $\mu_{\mathrm{DFA}}.\Psi_k$ the rate (density) of deforestation in the network or possibly networks (if $m > 1$) intersected by the plot. In SRS, it is the DFA per unit plot area within each plot that is recorded. Hence, the variation in DFA among PUs within a deforestation network does not contribute to the sampling variance in ACS as it does in SRS. The variance efficiency of an ACS sampling design therefore depends on the magnitude of the within-network variance $\sigma^2_{\mu_{\mathrm{DFA}}}$ (within Ψ) relative to that of the among-network variance $\sigma^2_{\mu_{\mathrm{DFA}}}$ (among Ψ) (Thompson 1992). We obtained population-specific values of these two variances to gauge the expected relative efficiency of the ACS designs (Table 2). The within-network variance of deforestation rates (density) was in most populations about one half the among-network variance but less than one tenth the among-network variance in three populations (average 0.30). These ratios play an important role for the variance efficiency of an ACS design relative to that of SRS.

Expected ACS sample sizes

In ACS, the adaptive expansion of the initial n SRS plots means that the number of sampled PUs will be larger than the nominal $n \times m$ units. Let $n \times m^*$ denote the number of PUs actually sampled under an ACS design with an initial sample size of n plots of size m PUs. The realized ACS sample size ν in plot equivalents is $n \times m^* \times m^{-1}$. It follows that ν is a design-dependent random variable that will vary between repeat sampling a population. In a known population, the expected value $E(\nu)$ of ν is (Thompson 1990)

$$E(\nu|m) = \sum_{i=1}^{M}\sum_{j=1}^{m} 1 - \binom{M - x_{ij} - a_{ij}}{n} \times \binom{M}{n}, \tag{10}$$

where x_{ij} is the number of plots of size m PUs intersecting the network containing the jth PU in the ith plot and a_{ij} is the number of PUs in networks with deforestation that adjoins the jth PU in the ith plot. Hence, a_{ij} accounts for the adaptive inclusion of edge units, a_{ij} is

Table 2 Among- $\sigma^2_{\mu_{DFA}}$ (among Ψ) and within-network variance of deforestation rates $\sigma^2_{\mu_{DFA}}$ (within Ψ). Columns 4–7 are the expected values of indicators of ACS efficiency. See text for definitions of indicators

POP	σ^2_{DFA} (plot$\mid m = 1$) $\times 10^2$	σ^2_{DFA} ($w \mid m = 1$) $\times 10^2$	$E(\nu)/n$	λ_{max}	E_ν	C_{max}
1	0.162	0.070	1.2	1.8	0.7	3.8
2	0.315	0.137	1.5	1.8	0.8	1.7
3	0.614	0.276	1.8	1.8	1.0	1.0
4	0.174	0.053	1.2	1.4	0.8	2.3
5	0.350	0.106	1.9	1.4	1.4	0.5
6	0.698	0.215	2.2	1.5	1.5	0.4
7	0.128	0.056	1.6	1.8	0.9	1.2
8	0.253	0.109	2.3	1.8	1.3	0.6
9	0.086	0.042	2.4	1.9	1.2	0.7
10	0.160	0.051	2.1	1.5	1.5	0.4
11	0.334	0.115	3.0	1.5	2.0	0.3
12	0.633	0.219	3.7	1.5	2.5	0.2
13	0.196	0.009	1.6	1.0	1.6	0.1
14	0.390	0.021	3.1	1.1	3.2	0.1
15	0.768	0.076	6.4	1.1	6.3	0.1
16	0.171	0.060	1.1	1.5	0.7	4.8
17	0.340	0.120	1.3	1.5	0.8	1.9
18	0.680	0.243	1.5	1.6	0.9	1.2
Mean	0.358	0.109	2.2	1.5	1.6	1.2

zero if the jth PU in the ith plot is not located next to any deforestation network.

Efficiency of ACS relative to SRS

For any given initial sample size n, plot size m, and population, the sampling variance of an ACS estimate of deforestation rate (density) will be less than the variance of the corresponding SRS estimate simply because more PUs are sampled with an ACS design. Implementing the adaptive sampling procedure incurs costs that are not encountered in SRS. The magnitude and structure of these costs will vary depending on the actual context, technology, and data capture procedures. A realistic comparison of ACS and SRS must consider the impact of the adaptive sampling on both cost and realized variance. Within an image analysis framework, however, the additional resources employed for the adaptive component of the ACS design may be reasonable.

The potential variance efficiency of an ACS design can be gauged by the maximum inflation λ_{max} of $n \times m$ that can be allowed before a SRS sampling of $\lambda_{max} \times n \times m$ PUs would yield the same expected sampling variance as the ACS design. Thompson (1992, p. 275) provides details for calculating λ_{max} which involve estimators of the within- and among-network variance of deforestation rates (density). For a constant cost of sampling one PU, ACS is only attractive if $\lambda_{max} \times n > E(\nu \mid m)$.

The ratio E_ν of the expected sampling variance of an ACS estimate of deforestation rate to the expected sampling variance of an SRS estimate of deforestation rate when ν plot equivalents are sampled in both cases provides another measure of ACS efficiency. A value of E_ν below 1.0 indicates that ACS is more (variance) efficient than SRS for a constant sample size of ν plots. E_ν does not, however, consider that the total cost of sampling under the two designs with equal ν may differ.

A third measure of ACS efficiency, C_{max}, is the upper limit that can be paid by ACS for every m PUs added adaptively to the initial sample under the assumption of equal expected sampling variances and equal total sampling costs. Should actual costs exceed C_{max} then, for fixed total costs, ACS will be less (variance) efficient than SRS. We use the relative cost ratio C_{max}/C where C is the cost of sampling m PUs under a SRS design as the benchmark of this efficiency. This relative cost ratio is not influenced by the value of C, for convenience we chose $C = 1.0$.

Auxiliary statistics for prediction of ACS performance

The expected relationship between population structure and the efficiency of ACS makes it desirable to have models that can predict the expected performance of ACS relative to that of SRS from a set of readily available auxiliary statistics. A survey designer could then use these models to gauge whether ACS is an attractive design option or not. Auxiliary statistics could come from a pilot study of DFP in a couple of representative areas (Leckie et al. 2002).

Since all aspects of the DFPs in the 18 populations are known to us, we can explore the potential for predicting the relative performance of ACS from a set of predictors such as, for example, rate of deforestation, number, shape, and size of DFPs, nearest neighbor distances of DFPs, and indices of aggregation of PUs with deforestation. Several candidate predictors were significantly correlated with the expected ACS sampling variance and to the above three measures of ACS design efficiency. The most parsimonious yet promising set of predictors was the rate of deforestation μ_{DFA}, the average number of PUs in a deforestation network ($\bar{\bar{n}}_{\Psi_{DFA}}$), and the average perimeter-to-area ratio of DFP ($\bar{R}_{P:A}$). The latter was computed as $(1/16)P^2 \times A^{-1}$ where P is the perimeter of a DFP (unit: m) and A its area (unit: m^2). A perfectly square DFP would have a $R_{P:A}$ value of 1. DFPs with a "linear" shape have very large ratios. Population-specific averages of these predictors are in Table 1. It is intuitive that these three predictors would have an effect on the efficiency of ACS. We expect that the prediction models developed here will be useful for pre-survey evaluations of design-options.

Results

Bias

The three estimators of deforestation rates $\hat{\hat{\mu}}_{DFA}$, $\hat{\hat{\mu}}_{DFA.HH}$, and $\hat{\hat{\mu}}_{DFA.HT}$ are asymptotically design-unbiased. An average of 400 repeat samples will, however, only approximate the true fixed population value of μ_{DFA}. For the sake of the comparison between SRS and ACS, we consider the difference between the average of 400 estimates of deforestation rates and the true population value as bias. Edge effects caused by plots straddling the population boundary (Williams 2001b, Gregoire and Scott 2003) have not contributed to the apparent bias since each population could be perfectly tessellated by each of the used plot types. Modified HH and HT estimates of deforestation rates both had a consistent (across populations and sample sizes) bias of ~3% when sampling with plots of size one PU. Bias of SRS estimators, in contrast, declined, as expected, with sample size, from 4% for $n = 200$ to 0.5% for $n = 3{,}000$. Sampling with multi-unit plots ($m = 10$ and 40) changed the bias profiles. For ACS, bias was positive (≈2%) for the smallest sample sizes and negative (≈−4%) for the largest sample sizes. Bias of SRS estimates started out negative (≈−10%) for the smallest sample size and was near 2% for the largest sample size. No single bias was statistically significant under the design-consistent hypothesis of no bias ($P > 0.24$, t-test). We also confirmed that the bias was neither significantly correlated with any available population statistic nor with any of 18 spatial statistics characterizing a population.

Empirical and expected sampling error

The reliability of the statistical estimators of sampling error was assessed by a comparison to the observed standard deviation of sample-based estimates in repeat sampling. For $m = 1$ and $n = 800$, the standard deviation of estimates $\hat{\hat{\mu}}_{DF.HH}$ and $\hat{\hat{\mu}}_{DF.HT}$ in 400 repeat samples matched the expected average of the modified HH and HT estimates of sampling error to within 4%. The fit improved with increasing n. For $n = 3{,}000$, discrepancies were less than 1%. Both ACS estimators of sampling error were, for $n \leq 400$, 10–15% lower than their empirical counterparts ($P > 0.19$, F_{max} test). In five populations (1, 4, 13, 14, and 16) and $n \leq 400$, the averages of the modified estimators were significantly below the empirical values. Similar problems were encountered with SRS estimators. Modified HH and HT estimators were, as a rule, within 2% of each other and strongly correlated (average $\rho = 0.98$). Results for plot sizes of 10 and 40 PUs were similar.

Root mean square errors

Modified HH and HT estimators of ACS relative mean square errors (RMSE%) were practically indistinguishable (differences always <1.7%). For the sake of brevity, we restrict our reporting of the main results to those obtained with the HT estimators.

RMSE% declined as a power function of sample size (RMSE% $= a \times n^{-b}$, $r^2 > 0.98$) with population-specific rates of decline (b) and "intrinsic" value (a). Population effects on RMSE% were significant for all test designs ($P < 0.01$, analysis of variance F-ratio tests). Population rankings of RMSE% across sample sizes were only moderately strong (coefficient of rank concordance (Moroney 1951) was 0.35 for $m = 1$, and about 0.7 for $m = 10$ and 40). Generally, a population with a high RMSE% for low n would improve its rank score as sample sizes increased and vice versa. An average correlation of the trend coefficients a and b of 0.60 supports this observation.

For ACS with $n < 800$ and $m = 1$, the median population RMSE% was above 40 (Fig. 4). Increasing n to 3,000, lowered the median RMSE% to 20. The RMSE% interval for the 16 centermost populations runs about 50% above and 50% below the median trend line. The median population RMSE% for SRS was 1.4 ($n = 200$) to 1.9 ($n = 3{,}000$) times higher than the corresponding ACS value. The improved performance of ACS for larger sample sizes was expected. As sample size increases, the probability of intersecting a deforestation network increases which translates into a marginal reduction of the within-network sampling variance—a reduction that is not possible under SRS. The correlation between SRS and ACS estimators of RMSE% was consistently strong and positive ($\rho > 0.98$) and the relationship was linear.

Designs with $m = 10$ and 40 confirmed the general trends for $m = 1$ except for a noticeable increase in the width of ~90% confidence intervals. The performance gap between the worst and the best in the group of 16 centermost populations was roughly twice as large when m was 10 and 40 compared to the results for $m = 1$. An increase in the relative variability of population-specific results for $m = 10$ and 40 reflects the fact that a concentration of sampling to fewer locations with larger

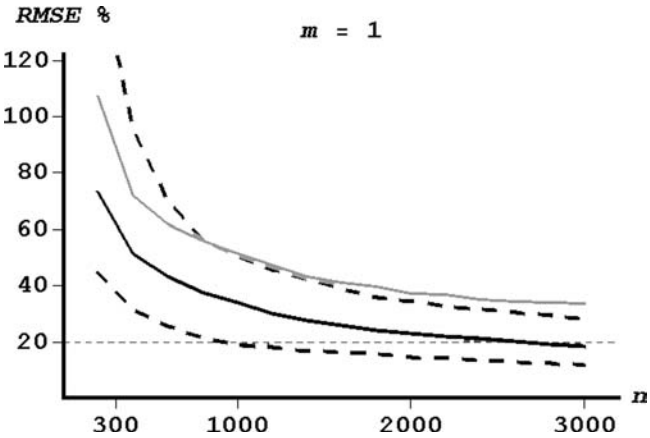

Fig. 4 Relative root mean square error (RMSE%) for ACS designs versus sample size (n). *Full black line*: median trend line; *dashed lines* are the upper and lower approx. 10% population quantile, respectively. *Full gray line*: median trend line for the SRS designs

plots is sensitive to the number, size, and spatial distribution of DFPs. The sample yield in terms of DFA will tend to vary as a function of the inverse of plot size. The expected lower (variance) efficiency of multi-unit plots (Magnussen 2001) in a design with a fixed number $n \times m$ of sampled PUs confirmed this expectation by a general inflation of RMSE% by a factor of about 1.4 ± 0.2 for $m = 10$ and 1.6 ± 0.3 for $m = 40$.

Efficiency of ACS

ACS achieved a 30–50% lower RMSE% by increasing, beyond the nominal $n \times m$, the number of sampled PUs. The ratio $E(\nu) \times n^{-1}$ of the number of plot equivalents sampled under ACS to that sampled under SRS was, on average for the 18 populations, 2.2 ($m = 1$) but varied from a low of 1.1 to a high of 6.4 (Table 2). Note that the expected value of $E(\nu)/n$ depends on n but within the range of n/N-values considered in this study the changes were minimal (<2%) and could safely be ignored. Highest ratios were seen in populations with a relatively high deforestation rate, DFP with a high perimeter-to-area ratio, and a strongly skewed size distribution of polygon deforestation areas. Individual population results for $m = 10$ and 40 were highly correlated with those for $m = 1$ (0.98 and 0.91), but with averages 23% and 27% lower, respectively.

The maximum ACS inflation $\lambda_{\max}$ of $n \times m$ that would give an SRS design with a sample size of $\lambda_{\max} \times n \times m$ PUs, the same expected sampling variance as an ACS design with an initial SRS sample of $n \times m$ PUs, averaged 1.5 across populations (Table 2). Since the average of $E(\nu) \times n^{-1}$ was 2.2 and in ten populations larger than $\lambda_{\max}$ it follows that the cost of sampling the extra PUs becomes a pivotal criterion for accepting ACS as a viable alternative. Only in seven populations was $E(\nu) \times n^{-1}$ below $\lambda_{\max}$ ($m = 1$). The expected variance efficiency E_ν of ACS in column 6 of Table 2 mirrors these results. A direct consequence of these results is that an increase in SRS sample sizes by a factor of $\lambda_{\max}$—in order to match the expected realized sample sizes under ACS—can be expected to give a sampling variance that is about 1.6 times lower and only in seven cases higher than the corresponding ACS sampling variance. Results for $m = 10$ and 40 did not reveal any new trend except for an overall downward shift of 10–20% in the expected values.

To achieve a break-even of costs and sampling variances, the maximum price ACS can afford to pay for every m PUs adaptively added to the initial sample is, on average across the 18 populations, 1.2 times the price of adding m such units in SRS (see $C_{\max}$ in Table 2). In eight populations and $m = 1$, the break-even price for ACS is less than 0.5. The low break-even prices in three populations of max 0.1 serve as a warning of the serious implications ACS can have on costs. Results for $m = 10$ and 40 were almost identical (within 10%). The need to screen out the ACS option for this type of population is clear. Fortunately, this screening appears to be within reach.

Predictions of ACS performance from auxiliary statistics

The performance of ACS depends strongly on the number, size, shape, and spatial distribution of DFPs. It is therefore critical for a survey designer to have a reasonable prediction of the expected performance in a given population before adopting an ACS design. A stratification of regions suitable and not suitable for ACS would be the objective of this pre-survey analysis. The predictors of ACS performance should be easily available from pilot studies providing representative maps of DFPs. We present a set of prediction models derived from the 18 populations with the expectation that these models can be used in practical settings to predict the expected performance of ACS for the estimation of deforestation rates.

The expected ACS sampling variance for $m = 1$ could be predicted with a standard error of prediction of $0.00033 \times n^{-1}$ from the population level of deforestation μ_{DFA}, the average number of PUs in a deforestation network $\bar{\bar{n}}_{\Psi_{\mathrm{DFA}}}$, and the average perimeter-to-area ratio of the DFPs in a population ($\bar{\bar{R}}_{P:A}$).

$$\begin{aligned}\sigma^2_{\mathrm{DFA.HT}}|m = &\left(0.00061 + \mu_{\mathrm{DFA}}^{1.548} \times \bar{\bar{n}}_{\Psi_{\mathrm{DFA}}}^{0.630} \times \bar{\bar{R}}_{P:A}^{-0.291}\right) \\ &\times n^{-1}, \\ R^2_{\mathrm{adj}} = 0.96.&\end{aligned} \tag{11}$$

Predictive models for $m = 10$ and 40 (not shown) allowed predictions with errors of $0.03 \times n^{-1}$ ($R^2_{\mathrm{adj}} = 0.95$) and $0.14 \times n^{-1}$ ($R^2_{\mathrm{adj}} = 0.99$), respectively.

The expected ACS sample size ν for $m = 1$ could be predicted with a standard error of prediction of $0.22 \times n^{-1}$ from the following equation

$$\widehat{E}(v|m=1) = \left(1 + \mu_{\mathrm{DFA}}^{0.484} \times \bar{\bar{n}}_{\Psi_{\mathrm{DFA}}}^{0.376} \times \bar{\bar{R}}_{P:A}^{0.495}\right) \times n, \quad R^2_{\mathrm{adj}} = 0.97. \tag{12}$$

Equations for $m = 10$ and 40 (not shown) obtained equally good fit and their coefficients were within 10% of those for $m = 1$. The expected variance efficiency and other efficiency measures can be obtained via these predictions and the within-network variance of deforestation areas in Table 2.

Discussion

ACS was introduced as a potential efficient design for estimation of rates, densities, or totals of rare population attributes with an aggregated spatial distribution (Thompson 1990, 1991a, 1991b). Despite evidence that ACS can achieve tangible variance efficiency gains, it is also clear that sampling for estimation of rates of deforestation with any reasonable precision (<20%) will demand large sample sizes and costs. Under favorable circumstances, ACS can be very efficient in terms of precision and cost. Practical experience with ACS remains scant and the results so far indicate that both design effects and costs can be highly variable (Lo et al. 1997; Brown and Manly 1998; Ringvall 2000; Christman 2002; Hanselman et al. 2003; Smith et al. 2003; Brown 2003). Modifications of ACS aimed at capping the number of units sampled, by, for example, introducing a stopping rule or by optimizing the adaptive protocol, do not appear to have achieved gains of practical importance (Mohammad and Seber 1997; Lo et al. 1997; Christman and Lan 2001; Su and Quinn 2003). Fortunately, the performance can be gauged from prior knowledge or pilot studies exploring the population structure in terms of size, number, shape, and distribution of patches (polygons) displaying the attribute of interest.

ACS is akin to unequal probability sampling of networks. Its efficiency depends critically on the correlation between the inclusion probability and the network attribute of interest (Brewer and Hanif 1983; Godambe and Thompson 1988; Magnussen 2002)—a correlation modified by network shape and size. In the context of DFP, the correlation was relatively strong in populations dominated by polygons with an outline resembling a square and relatively weak in populations dominated by linear polygons of roads, hydro-lines, and similar man-made structures. Yet the effect of population structure on ACS efficiency remains complex. For example, the mean perimeter-to-area ratio of the DFP influences the efficiency of ACS. A low ratio decreases the sampling variance but also increases the expected sample size. The non-linear power functions predicting the performance of ACS highlighted interdependencies among the predictors and hence the need to explore carefully the population structure before committing to ACS. The cost of adaptively adding a PU to the sample is likely to be the overriding critical factor for the practical acceptance of ACS. Our results make it clear that a "window of opportunity" for ACS clearly exists in some populations but also that it is marginal in others. A sizeable efficiency gain was only manifest in seven of eighteen populations.

Forestry applications of ACS are scarce. Acharya et al. (2000) applied ACS for the estimation of abundance of three rare tree species in a 40-ha compact region in Nepal. The region was subdivided into a grid of square (5 m × 5 m) PUs. ACS was highly efficient compared to SRS for one species yet inefficient for the other two. The species-specific performance of ACS was attributed to species-specific differences in the spatial distribution of presence/absence. Roesch (1993) developed estimators for a point adaptive sampling design. They were tested in a trial with eight rare tree species growing within a 3.1-ha stem-mapped forest stand in Maine. Simulated sampling indicated that ACS was often more efficient than SRS, but it was not uniformly superior across all eight species.

Although our study is based on simulated sampling from artificial populations, the effort that went into creating 18 realistic populations of DFPs and the nontrivial size of a population vouch for the practical relevance of our results. The sampling designs were limited to sample sizes and plot sizes deemed sufficient for the exploration of the performance of ACS relative to that of SRS. Our sample populations reflect the expected range of regional differences. It has been demonstrated that for the same number of PUs in a plot, a long and narrow plot will be relatively more efficient than a square plot (Magnussen 2001). While a multi-unit plot design is less (variance) efficient per PU of observation than a single-unit plot design (Cochran 1977; Magnussen et al. 1998), the concentration of the sampling effort to fewer but larger plots may still be advantageous from both a logistical and a cost perspective. Finding the most suitable sampling design for estimation of deforestation rates in Canada was beyond the focus of this study. Our equations predicting the expected RMSE% for a given sample size should give the survey planner a reasonable prediction for designs outside the study range. The search for a favorable national design is ongoing. Issues and factors to consider at a national scale extend well beyond what we have covered.

Intuitive models predicting the expected ACS sampling variance, and the expected realized ACS sample size confirmed a significant effect of deforestation rate, the size of deforestation networks, and the shape of the DFP on the performance of ACS. The real significance of these models is that they allow a survey designer to obtain a reasonable prior expectation about the performance of ACS whenever the surveyor can get hold of reasonable estimates of the predictors—information that is also useful for stratifying a population. Regional differences in the characteristics of the predominant DFP favor stratification. The recognized sensitivity of ACS efficiency to population structure has been a serious impediment to wider acceptance (Roesch 1993; Lo et al. 1997; Acharya et al. 2000; Hanselman et al. 2003; Smith et al. 2003). Published ACS studies do not appear to have attempted to relate the performance of ACS to various descriptive population statistics.

Without a stopping rule, ACS runs a non-trivial risk of serious cost overruns, a phenomenon well known to practitioners of sampling with probability proportional to predictions (Bonnor 1972; Mohammad and Seber 1997; Lo et al. 1997; Brown and Manly 1998; Christman and Lan 2001; Muttlak and Khan 2002; Su and Quinn 2003). ACS is seemingly unattractive in populations with a relative low number of DFP dominated by a few and very large polygons unless the marginal cost of adding a large number of PUs to the sample is low. The characteristics of populations unsuitable for ACS ought to make them relatively easy to screen out for a different sampling design during a pre-stratification process.

Application of ACS to ecological field surveys has been limited by the requisite subdivision of a population into PUs of equal size and shape. The layout and ACS sampling of units could be a costly and logistically challenging endeavor (Roesch 1993; Acharya et al. 2000; Christman 2000, 2002; Hanselman et al. 2003; Smith et al. 2003). A recent extension of ACS to line intercept sampling (Morgan 1997) and the aforementioned extension to point sampling (Roesch 1993) alleviate this problem to a degree. ACS sampling within a framework of image analysis could, in principle, be fairly easy to implement since image pixels would be the natural sampling frame to work from. The costs incurred by the adaptive protocol of ACS may in fact be low, especially if the image analysis for the mapping of DFP is object oriented (Coppin et al. 2004).

PUs are either defined naturally by the context, or chosen by the survey designer. In the latter case, the size of the unit will be important for the cost structure and possibly the efficiency of ACS. With either extremely small or large units, relative to the size of DFP, ACS cannot be more efficient than SRS since the within-network unit variance of deforestation areas vanishes at the extremes. This happens at the high end because most deforestation networks will fit entirely inside a unit and at the low end because a vast majority of units in a deforestation network will have a deforestation area equal to the size of the unit. We examined the impact of unit size by computing the among- and the within-network variance of deforestation areas for units of size 50 $\times$ 50 m^2 (0.25 ha), 150 $\times$ 150 m^2 (2.25 ha), and 200 $\times$ 200 m^2 (4 ha). These calculations were done on a subset of 12 populations and with at least 10% of the DFPs in a selected population. Although preliminary in nature, the results did not indicate any marked deviations in the relative performances of ACS and SRS seen with the 1 ha PU. Expected sample size ratios stayed within 20% of the results with the 1-ha unit without any discernable trend across unit size. The same holds for the ratios of SRS to ACS sampling variances. While the unit choice of 1 ha was somewhat arbitrary, the choice seems inconsequential within the above range of sizes.

Logical alternatives to ACS are sampling based on model predictions (Ståhl et al. 2000; Williams 2001a) and possibly ranked subset sampling in which cheap but error-contaminated samples are collected first and ranked for deforestation area, a much smaller error-free but expensive sample is then taken in order to calibrate the rankings (Barnett 1999; Chen 1999; Nabhas et al. 2002). The opportunity to expand an existing survey into an adaptive one, would, everything else being equal, favor ACS.

Both the modified HH and the modified HT estimators of deforestation rates and sampling errors performed well in terms of bias and their match to empirical estimates of standard error. Modified HT estimators have been reported to be less sensitive to population structure (Su and Quinn 2003; Felix-Medina 2003; Salehi 2003), an observation that was only partially confirmed by this study. HT estimates were only slightly less biased and less variable than HH estimates, and differences were not considered as important. The apparent complexity of the HT variance estimators may be the decisive factor. By including both, however, one also obtains an efficient check against programming errors. A Rao–Blackwell reduction of the HT/HH sampling variance estimates (Thompson 1991a; Salehi 1999; Mohammad 2002) was not pursued here due to the large number of deforestation networks (> 1,000 in most populations) which would make computations onerous.

Natural resource surveys are rarely designed for a single attribute (Schreuder et al. 1993; Shiver and Borders 1996). Rather, the tendency is for surveys to collect a wide array of tree, vegetation, and soil attributes within a single survey. Sampling designs that are optimized for sampling a rare attribute may be inefficient for other attributes (Roesch 1993; Green 1993; Acharya et al. 2000; Christman 2000; Dryver 2003). Thus, natural resource designs tend to be compromises of several conflicting demands (Nusser et al. 1998; Olsen et al. 1999; Corona et al. 2002). Rare attributes are therefore usually estimated relatively poorly by multi-purpose surveys (Green 1993; Christman 2002; Venette et al. 2002). When it becomes important to have reliable estimates of a rare and spatially aggregated attribute, the best sampling strategy may be to combine the results from an existing multi-purpose survey with auxiliary sampling (Samuel-Cahn 1994; Peña 1997). Fortunately, the ACS design is flexible enough to permit integration with an already existing survey (Roesch 1993).

The general conclusion of this study is that the 30–50% lower RMSE achieved with ACS remains attractive when the cost of obtaining each extra unit added to the sample is not higher than the cost of sampling a unit under SRS. Practical considerations will determine plot size. ACS deserves careful consideration along with more traditional options when sampling for deforestation rates and other rare but spatially aggregated population attributes.

Acknowledgments Financial support for this study was provided, in part, by the Government of Canada's Climate Change Action Fund. We are indebted to Mrs. Ruth Parnall, Canadian Forest Service, Edmonton for her help and diligence in creating the 18 populations of DFP.

References

Acharya B, Bhattarai G, de Gier A, Stein A (2000) Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. Forest Ecol Manage 137:65–73

Barnett V (1999) Ranked set sample design for environmental investigations. Environ Ecol Stat 6:59–74

Bonnor GM (1972) A test of 3-P sampling in forest inventories. Forest Sci 18:198–202

Brewer KRW, Hanif M (1983) Sampling with unequal probability. Lecture notes in statistics, vol 15. Springer, New York, p 164

Brown JA (2003) Designing an efficient adaptive cluster sample. Environ Ecol Stat 10:95–105

Brown JA, Manly BJF (1998) Restricted adaptive cluster sampling. Environ Ecol Stat 5:49–63

Chen ZH (1999) Density estimation using ranked-set sampling data. Environ Ecol Stat 6:135–146

Christman MC (2000) A review of quadrat-based sampling of rare, geographically clustered populations. JABES 5:168–201

Christman MC (2002) Efficiency of some sampling designs for spatially clustered populations. Envirometrics 8:145–166

Christman MC, Lan F (2001) Inverse adaptive cluster sampling. Biometrics 57:1096–1105

Cochran WG (1977) Sampling techniques. Wiley, New York

Coppin PR, Jonckheere I, Nackaerts K, Muys B, Lambin EF (2004) Digital change detection methods in ecosystem monitoring: a review. Int J Rem Sens 25:1565–1596

Corona P, Chirici G, Marchetti M (2002) Forest ecosystem inventory and monitoring as a framework for terrestrial natural renewable resource survey programmes. Plant Biosyst 136:69–82

Correll RL, Cellier KM (1987) Effects of plot size, block size and buffer rows on the precision of forestry trials. Aust Forest Res 17:11–18

Dessai S, Schipper EL (2003) The Marrakech Accords to the Kyoto Protocol: analysis and future prospects. Global Environ Change Hum Policy Dimen 13:149–153

Dryver AL (2003) Performance of adaptive cluster sampling estimators in a multivariate setting. Environ Ecol Stat 10:107–113

Evelyn OB, Camirand R (2003) Forest cover and deforestation in Jamaica: an analysis of forest cover estimates over time. Int Forestry Rev 5:354–363

Felix-Medina MH (2003) Asymptotics in adaptive cluster sampling. Environ Ecol Stat 10:61–82

Fitzsimmons M (2002) Estimated rates of deforestation in two boreal landscapes in central Saskatchewan, Canada. Can J Forest Res 32:843–851

Fuller DO, Jessup TC, Salim A (2004) Loss of forest cover in Kalimantan, Indonesia, since the 1997–1998 El Nino. Conserv Biol 18:249–254

Godambe VP, Thompson ME (1988) On single stage unequal probability sampling. In: Krishnaiah PR, Rao CR (eds) Sampling. Elsevier, Amsterdam, pp 111–123

Green RH (1993) Sampling to detect rare species. Ecol Appl 3:351–356

Gregoire TG, Scott CT (2003) Altered selection probabilities caused by avoiding the edge in field surveys. J Agric Biol Ecol Stat 8:36–47

Hanselman DH, Quinn TJ, Lunsford C, Heifetz J, Clausen D (2003) Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bull 101:501–513

Koop G, Tole L (2001) Deforestation, distribution and development. Global Environ Change Hum Policy Dimen 11:193–202

Leaf D, Verolme HJH, Hunt WF (2003) Overview of regulatory/policy/economic issues related to carbon dioxide. Environ Int 29:303–310

Leckie DG, Gillis MD, Wulder MA (2002) Deforestation estimation for Canada under the Kyoto Protocol: a design study. Can J Rem Sens 28:672–678

Levy PE, Milne R (2004) Estimation of deforestation rates in Great Britain. Forestry 77:10–16

Lo NCH, Griffith D, Hunter JR (1997) Using a restricted adaptive cluster sampling to estimate Pacific hake larval abundance. Calif Coop Oceanic Fisheries Invest Rep 38:103–113

Madden LV, Hughes G (1999) Sampling for plant disease incidence. Phytopathology 89:1088–1103

Magnussen S (2001) Fast pre-survey computation of the mean spatial autocorrelation in large plots composed of a regular array of secondary sampling units. Math Model Sci Comp 13:204–217

Magnussen S (2002) Evaluation of probability proportional to predictions estimators of total stem volume. Can J Forest Res 32:92–102

Magnussen S, Boudewyn P, Gillis MD (1998) Towards a plot size for Canada's national forest inventory. In: Hansen M, Burk TE (eds) Integrated tools for natural resources inventories in the 21st century, 16 August 1998. USDA Forest Service, St. Paul, MI, pp 116–128

Mohammad SM (2002) Rao–Blackwell versions of the Horvitz–Thompson and Hansen–Hurwitz in adaptive cluster sampling. Environ Ecol Stat 6:183–195

Mohammad SM, Seber GAF (1997) Two-stage adaptive cluster sampling. Biometrics 53:970

Morgan BJT (1997) Strip adaptive cluster sampling: probability proportional to size selection of primary units. Biometrics 53:1096

Moroney MJ (1951) Facts from figures, 2nd edn. Penguin, London

Muttlak HA, Khan A (2002) Adjusted two-stage adaptive cluster sampling. Environ Ecol Stat 9:111–120

Nabhas RW, Wolfe DA, Chen H (2002) Ranked set sampling: cost and optimal set size. Biometrics 58:964–971

Nusser SM, Breidt EJ, Fuller WA (1998) Design and estimation for investigating the dynamics of natural resources. Ecol Appl 8:234–245

Olsen AR, Sedransk J, Edwards D, Gotway CA, Liggett W, Rathburn S, Reckhow KH, Young LJ (1999) Statistical issues for monitoring ecological and natural resources in the United States. Environ Monit Assess 54:1–45

Peña D (1997) Combining information in statistical modeling. Am Stat 51:326–332

Pollard JH, Palka D, Buckland ST (2002) Adaptive line transect sampling. Biometrics 58:862–870

Ringvall A (2000) Assessment of sparse population in forest inventory: development and evaluation of probability sampling methods. PhD Thesis, Swedish University of Agricultural Sciences, Umeå

Roesch FAJ (1993) Adaptive cluster sampling for forest inventories. Forest Sci 39:655–669

Sader SA, Hayes DJ, Hepinstall JA, Coan M, Soza C (2001) Forest change monitoring of a remote biosphere reserve. Int J Rem Sens 22:1937–1950

Salehi M (1999) Rao–Blackwell versions of the Horvitz–Thompson and Hansen–Hurwitz in adaptive cluster sampling. Environ Ecol Stat 6:183–195

Salehi M (2003) Comparison between Hansen–Hurwitz and Horvitz–Thompson estimators for adaptive cluster sampling. Environ Ecol Stat 10:115–127

Samuel-Cahn E (1994) Combining unbiased estimators. Am Stat 48:34–36

Särndal C-E, Swensson B, Wretman J (1992) Model assisted survey sampling. Springer, New York
Schreuder HT, Gregoire TG, Wood GB (1993) Sampling methods for multiresource forest inventory. Wiley, New York
Shiver BD, Borders BE (1996) Sampling techniques for forest resource inventory. Wiley, New York
Smith DR, Villella RF, Lemarie DP (2003) Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environ Ecol Stat 10:7–15
Srivastava S, Singh TP, Singh H, Kushwaha SPS, Roy PS (2002) Assessment of large-scale deforestation in Sonitpur district of Assam. Curr Sci 82:1479–1484
Ståhl G, Ringvall A, Lamas T (2000) Guided transect sampling for assessing sparse populations. Forest Sci 46:108–115
Stier SC, Siebert SF (2003) Tropical reforestation and deforestation and the Kyoto Protocol. Conserv Biol 17:5–6
Su ZM, Quinn TJ (2003) Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environ Ecol Stat 10:17–41
Thompson SK (1990) Adaptive cluster sampling. J Am Stat Assoc 85:1050–1059
Thompson SK (1991a) Adaptive cluster sampling: designs with primary and secondary units. Biometrics 47:1103–1115
Thompson SK (1991b) Stratified adaptive cluster sampling. Biometrika 78:389–397
Thompson SK (1992) Sampling. Wiley, New York
Upton GJG, Fingleton B (1985) Spatial data analysis by example. Vol I. Point patterns and quantitative data. Wiley, Chichester
Venette RC, Moon RD, Hutchison WD (2002) Strategies and statistics of sampling for rare individuals. Ann Rev Entomol 47:143–174
Williams MS (2001a) Non-uniform random sampling: an alternative method of variance reduction for forest surveys. Can J Forest Res 31:2080–2088
Williams MS (2001b) Comparison of estimation techniques for a forest inventory in which double sampling for stratification is used. For Sci 47:563–576