Environ Monit Assess (2016) 188: 11
DOI 10.1007/s10661-015-5002-8

Precision and cost considerations for two-stage sampling in a panelized forest inventory design

James A. Westfall · Andrew J. Lister · Charles T. Scott

Received: 13 July 2015 / Accepted: 20 November 2015 / Published online: 4 December 2015

Abstract Due to the relatively high cost of measuring sample plots in forest inventories, considerable attention is given to sampling and plot designs during the forest inventory planning phase. A two-stage design can be efficient from a field work perspective as spatially proximate plots are grouped into work zones. A comparison between subsampling with units of unequal size (SUUS) and a simple random sample (SRS) design in a panelized framework assessed the statistical and economic implications of using the SUUS design for a case study in the Northeastern USA. The sampling errors for estimates of forest land area and biomass were approximately 1.5–2.2 times larger with SUUS prior to completion of the inventory cycle. Considerable sampling error reductions were realized by using the zones within a post-stratified sampling paradigm; however, post-stratification of plots in the SRS design always provided smaller sampling errors in comparison. Cost differences between the two designs indicated the SUUS design could reduce the field work expense by 2–7 %. The results also suggest the SUUS design may provide substantial economic advantage for tropical forest inventories, where remote areas, poor access, and lower wages are typically encountered.

J. A. Westfall (✉) · A. J. Lister · C. T. Scott
U.S. Forest Service Northern Research Station, Newtown Square, PA, USA
e-mail: jameswestfall@fs.fed.us

A. J. Lister
e-mail: alister@fs.fed.us

C. T. Scott
e-mail: ctscott@fs.fed.us

Keywords Stratification · Subsampling · Finite population · Fieldwork efficiency · Inventory logistics

Introduction

Numerous sampling designs have been employed to inventory and monitor forests (Tomppo et al. 2010). Well-planned designs consider various technical requirements to achieve the defined inventory objectives and information needs, including data sources, plot design, sample size, use of ancillary data, and available funding. Ultimately, the goal is to obtain forest resource estimates having acceptable precision for minimal cost (Köhl et al. 2011). These factors, along with other considerations such as data currency and logistical constraints, often differ depending on the geographic location and spatial extent of the inventory effort. From a national forest inventory (NFI) perspective, many nations have implemented systematic sampling designs across the entire population (Kändler 2009; Brassel and Lischke 2001). In various countries, there has been a tendency to shift away from periodic inventories in favor of panel-type designs where a portion (often 10–20 %) of the plots are measured each year, with remeasurement of the same plots occurring at pre-specified intervals (e.g., 5–10 years) (Gillis et al. 2005; McRoberts et al. 2010). Plots selected to be measured in a given year are often distributed across the population in a spatially balanced (interpenetrating) manner. This leads to similar, relatively large inter-plot distances for a given panel, as nearby plots are scheduled for measurement in other years. This approach is clearly suboptimal in regard to field work efficiency and can be particularly problematic when areas with limited accessibility or difficult terrain need to be visited on a recurring basis.

Often, the primary limiting factor in obtaining estimates of sufficient precision is the amount of financial resources available; this creates an emphasis on obtaining the most precise estimates for a fixed cost. Measuring sample plots in the field is necessary to obtain observations of many important forest characteristics; however, this effort is almost always the most costly component of forest inventory operations, and substantial effort is often employed to optimize field work efficiency via appropriate specification of the sample and plot designs, data collection protocols, and sample size (Köhl and Scott 1992). To this end, an often overlooked design to assess forest resources is subsampling with units of unequal size (Cochran 1977, sec. 11.7). also referred to as a form of cluster sampling (Kangas and Maltamo 2006, p. 28) or two-stage sampling (Loetsch 1957; Gregoire and Valentine 2008, sec. 12). Under this paradigm, spatially distinct zones (e.g., compartments, management areas, or areas readily accessed in a single trip) are the primary sampling units (PSU), and plots within each zone are secondary sampling units (SSU).[1] In forest inventories, zones are usually not of equal size, thus the subsampling with units of unequal size (SUUS) terminology. Operationally, all plots in a given zone are completed prior to advancing to the next selected zone. This makes field work more efficient since less time is spent traveling between plots compared to designs that do not employ clustering. Furthermore, zone-level estimates may be desirable (such as for a compartment or management area) and can be obtained immediately after all the plots in a zone have been measured.

Because of the similarity of plots within zones, the variance of a two-stage sample will be higher than that of a simple random sample (SRS) with the same number of plots. Thus, although complete sample information for some zones is acquired in the initial years of implementation of such a design, it is at the expense of less precise estimates for the population. The precision improves each year as more zones are sampled. Upon completion of the inventory cycle, all selected zones and plots will have been measured. If every zone in the population is sampled (i.e., complete enumeration of all possible PSUs) and the probabilities of selection in the second stage are consistent across all zones, the secondary units (plots) may be treated as the PSUs and estimation may proceed under an SRS or stratified SRS paradigm (where zones act as strata). The choice of the SUUS design with complete enumeration of zones requires an evaluation of the balance between the field-work cost savings and temporarily higher standard errors of estimates. In this paper, we compare statistical estimates and inventory costs between SUUS and SRS paneled designs under both unstratified and stratified estimation scenarios with the goal of providing a framework for the evaluation of this cost/precision balance. These analyses will assist forest inventory planners in evaluating alternative cost/precision options and understanding the implications of selecting a SUUS design.

Methods

Data for this analysis was from completed inventory cycles over the period 2007 to 2011 in the US states of New York, Pennsylvania, and Ohio collected by the Forest Inventory and Analysis (FIA) program of the US Forest Service. The permanent sample plots are distributed across the landscape in a spatially balanced (systematic) manner, with one randomly located plot in each 2403-ha hexagonal grid cell with approximately 5 km between cell centers. This sample is treated as a simple random sample (SRS; Reams et al. 2005). Phase one (P1) of the FIA inventory design is to post-stratify the existing field plots using remotely sensed data in order to increase precision of the estimates (Bechtold and Patterson 2005). Geospatial data, typically derived from National Land Cover Database (NLCD) products (Homer et al. 2004). are used to create non-overlapping strata. Each plot is assigned to a stratum via spatial overlay, and the strata weights are calculated as the

[1] There are two common uses of the term "cluster sampling." The first is what we refer to as a "cluster plot," where the plot is not contiguous but is composed of two or more subplots separated in a specified nonrandom arrangement (Tokola and Shrestha 1999; Bechtold and Scott 2005). The second usage refers to a two-stage design in which zones (PSU) are randomly selected and plots within the zone are a random subsample of secondary sampling units (SSU) (Stehman et al. 2009). It is the latter usage that is adopted in this paper.

proportion of the population occupied by each stratum. The second phase (P2) entails measuring sample plots for the proportion of the plot that is forested and the usual suite of forest mensuration variables including tree species, diameter at breast height (d.b.h.), height, site index, forest type, and stand age. For each sample plot, trees with a d.b.h. of 12.7 cm or larger are measured on each of four 7.3-m radius subplots. In this study, standing woody biomass and area of forest land use on P2 plots were selected as the attributes of interest, since these are common variables of interest to forest inventory data users.

To simulate an SUUS design, a "population" of zones from which to sample was needed. The FIA plot data was divided into zones (PSUs) within which the plots (SSUs) were selected. In order to assess the effects of zone size on outcomes, four different fixed populations were produced having an average size (and range) of 5 (3–7), 10 (8–12), 15 (13–17), and 20 (18–22) plots per zone. The zone sample sizes were allowed to vary slightly around the average size to mimic practical solutions likely to be employed operationally. To group spatially proximate plots, a space filling curve was generated that passed through each plot in the population (Lister and Scott 2009). The space filling curve is a fractal, or repeating line pattern, that passes through every plot in the study area. This line can be broken into segments that will tend to cluster constituent points into spatially compact groups. For example, zones with an average of 10 plots were formed by dividing this line into sequential segments containing between 8 and 12 plots (Fig. 1). Polygonal areas were created to define each zone so that each element in the population was accounted for. Using a geographic information system (GIS) procedure, a grid of points with 1-km spacing was superimposed on the study area. These points were assigned to the nearest plot and subsequently to the zone in which the plot resided. In this way, each square kilometer of the population was assigned to a zone and zone boundaries were delineated. Due to the nearly uniform spatial distribution of the plots (Bechtold and Patterson 2005). the zones had approximately equal sample intensities (1 plot per 2403 ha). In practice, a SUUS design would be implemented by partitioning the population into zones (PSUs) and selecting a sample of zones based on a sampling design. Similarly, in the zones to be sampled, plots (SSU) would be selected under a specified design. In this case, the existing FIA spatially balanced sample grid was used, so plots were selected prior to the partitioning of the zones. A simulation study verified there was no practical effect on the estimated variances attributable to this difference in convention. Thus, while the SUUS design was not constructed in the usual manner, the properties of our conceptual formulation should be similar to those that would be found under a traditional implementation.

The SUUS design differs from the interpenetrating panel design FIA actually employed in these three

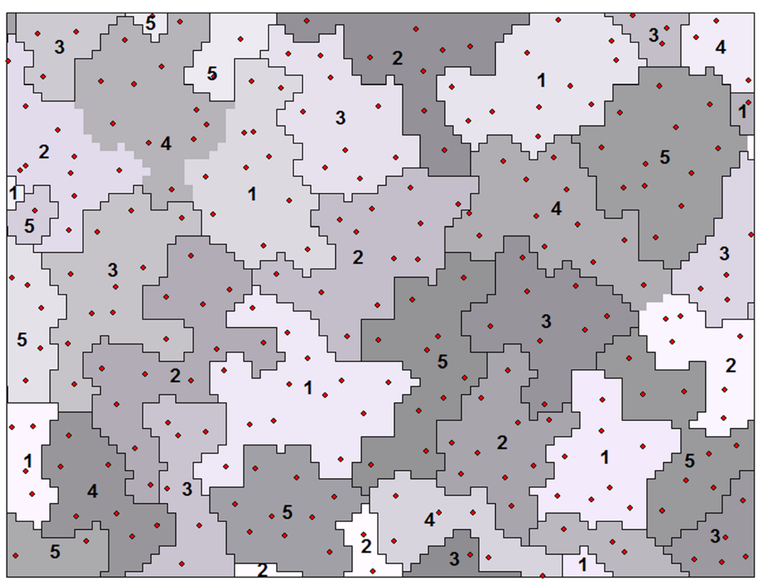

Fig. 1 Example of a panelized two-stage sampling design having an average of 10 plots (*diamond-shaped points*) in each zone. The zones were panelized into five spatially balanced subsets of the study area (as depicted by numerical assignments of 1–5); all zones in the same panel are measured in the same year

states, which entails measuring 20 % of the inventory plots across the population each year (a five-panel design). A similar number of plots would be measured each year under the SUUS framework by measuring 20 % of the zones each year. As is common practice for forest inventories where a portion of the plots are measured each year, temporal indifference was assumed; thus, zones or plots from different inventory years were combined to make estimates. Summaries of forest area proportion and biomass per hectare for zones and plots are provided in Table 1.

Estimation

Estimation for the SUUS design entails summing totals across sampled zones and expanding this value to the population level. The variance estimator includes both within- and between-zone components of uncertainty. The unbiased estimators for population totals in two-stage sampling with unequal size units are (Cochran 1977, Sec. 11.7):

$$\hat{Y}_{\mathrm{SUUS}} = \frac{N}{n}\sum_{i=1}^{n} M_i\overline{y}_i = \frac{N}{n}\sum_{i=1}^{n}\hat{Y}_i \tag{1}$$

$$V\left(\hat{Y}_{\mathrm{SUUS}}\right) = \frac{N^2(1-f_1)}{n}\frac{\sum_{i=1}^{n}\left(\hat{Y}_i-\widehat{\overline{Y}}\right)^2}{n-1} + \frac{N}{n}\sum_{i=1}^{n}\frac{M_i^2(1-f_2)s_i^2}{m_i} \tag{2}$$

where the notational definitions are given in Table 2. The above estimator was preferred over a ratio-to-size (Cochran 1977, Sec. 11.8) estimator due to the unbiased property; further, there were potential nonlinear relationships and generally low linear correlation between cluster size and the attributes of interest (forest area and biomass).

To assess the effects of the SUUS design on precision relative to an SRS design during the first cycle of implementation, a comparison was conducted by evaluating uncertainty in the estimates over a 5-year cycle wherein approximately 20 % of the sample was measured each year. For SRS, the existing FIA panel assignments for each plot were used to determine the plots to be measured in a given year. For SUUS, the corresponding effort would be the measurement of 20 % of the zones each year. Because there is no pre-specified assignment of zones to panels, it was desirable to assess the average outcome for the SUUS design over numerous potential zone assignment scenarios. This was accomplished by performing 1000 replications of randomly assigning zones to panels and calculating the estimates. The mean values of the estimates were calculated and reported in the results. These values were compared to those obtained under the SRS design. The SRS estimators for population totals are as follows:

$$\hat{Y}_{\mathrm{SRS}} = \sum_{i=1}^{N} M_i \frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i} y_{ij}}{\sum_{i=1}^{n} m_i} = \mathrm{M}\overline{\mathrm{y}} \tag{3}$$

$$V\left(\hat{Y}_{\mathrm{SRS}}\right) = M^2\frac{(1-f_2)\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left(y_{ij}-\overline{y}\right)^2}{\sum_{i=1}^{n} m_i\left(\sum_{i=1}^{n} m_i-1\right)} \tag{4}$$

where notational definitions are given in Table 2.

Table 1 Summary statistics for proportion forest and biomass (tons/ha) for plots (SRS) and four SUUS zone sizes

		Proportion forest				Biomass (tons/ha)			
Zone size	*n*	Min.	Mean	Max.	Std.dev.	Min.	Mean	Max.	Std.dev.
SRS	14,619	0.00	0.46	1.00	0.46	0	60.28	484.37	82.79
5	2937	0.00	0.46	1.00	0.31	0	60.48	327.22	51.84
10	1476	0.00	0.46	1.00	0.27	0	60.36	280.00	44.69
15	973	0.00	0.46	1.00	0.25	0	60.11	259.04	41.75
20	728	0.00	0.46	1.00	0.25	0	60.27	214.38	40.43

Estimators (1) and (2) are predicated on zones and the plots therein constituting an equal probability sample of the population; similarly, estimators (3) and (4) are appropriate for an equal probability sample of plots. However, the underlying FIA design distributes plots in a quasi-systematic manner via the hexagonal grid described above. As is commonly done in the absence of design-based variance formulae for systematic samples (Kangas and Maltamo 2006). estimators suited to simple random samples are employed with the knowledge that the variance estimate is often too large (Schreuder et al. 2004). The results are based on this common practice, and no comparisons were made to the unknown systematic sample variance.

As noted above, when the inventory cycle is completed for SUUS with complete measurement of all possible PSUs, all zones have been measured, and thus, there is no longer a sample of the primary units. At that time, the zones can be ignored and the data analyzed as a SRS design (assuming temporal indifference). Alternatively, the SUUS estimators could still be used if one considered the zones to now be strata, i.e., when all zones are measured, the SUUS estimators are equivalent to those for a stratified simple random sample design. However, the primary intent of stratification is to group observations having similar values such that a reduction in variance is achieved (Cochran 1977). A key advantage of SUUS is improved fieldwork efficiency derived from grouping plots based on spatial proximity rather than on grouping plots having similar attributes. Thus, these strata may be poorly defined for variance reduction. On the other hand, plots located near one another tend to exhibit autocorrelation of forest characteristics due to patterns in the environment; this autocorrelation might provide some stratification benefit.

Stratification of zones was also investigated as a way to incorporate stratification into the SUUS design. Theoretically, the form of the SUUS variance estimator (2) suggests that grouping of zones with similar population totals (rather than means) would provide strata conducive to reducing variance. Intuitively, zones having similar amounts of forested canopy cover may have similar amounts of forest land area and biomass. Using the NLDC land cover product (Homer et al. 2004). the number of map pixels with a canopy cover greater than 10 % was calculated for each zone. Strata were defined by various combinations of the quantiles of the distribution of these per zone values. To avoid small sample size issues and accommodate varying numbers of strata, these combinations were chosen to produce strata of approximately equal size.

To evaluate variance reduction in the context of post-stratifying zones, analyses were conducted for varying numbers of strata. Generally, increasing the number of strata results in greater precision, but at a decreasing rate. For each combination of number of strata and panels, estimates were generated by randomly assigning zones to panels for each of 1000 iterations. To compare effects of stratifying the plots treated as an SRS with those of the SUUS stratification, plots treated as an SRS were post-stratified using the same NCLD product (Homer et al. 2004). In this case, five strata were defined by percent canopy cover classes (0–5 %, 6–50 %, 51–65 %, 66–80 %, 81–100 %), and plots were assigned to strata based on the map-based canopy cover of the pixel containing plot center. Post-stratified estimation was accomplished by applying the SUUs (1, 2) and SRS (3, 4) estimators previously described within each stratum and then combining the estimated totals and their variances across all strata.

$$\hat{Y}^{ps}_{\text{SUUS}} = \sum_{h=1}^{H} \left(\hat{Y}_{\text{SUUS}}\right)_h \tag{5}$$

$$V\left(\hat{Y}^{ps}_{\text{SUUS}}\right) = \sum_{h=1}^{H} V\left(\hat{Y}_{\text{SUUS}}\right)_h \tag{6}$$

$$\hat{Y}^{ps}_{\text{SRS}} = \sum_{h=1}^{H} \left(\hat{Y}_{\text{SRS}}\right)_h \tag{7}$$

$$V\left(\hat{Y}^{ps}_{\text{SRS}}\right) = \sum_{h=1}^{H} V\left(\hat{Y}_{\text{SRS}}\right)_h \tag{8}$$

where notational definitions for Eqs. (5)–(8) are given in Table 2.

The variance estimators shown in (6) and (8) slightly underestimated the variances, as the stratification was employed after the sample units were selected, and the additional variance due to random samples sizes within strata was not accounted for. However, the estimators used are sufficient for comparative purposes as these estimators provide similar results to their post-stratified counterparts when stratum sample sizes are sufficiently

Table 2 Notational definitions for Eqs. (1)–(8)

$\hat{Y}$ = estimated population total
$\widehat{\overline{Y}}$ = estimated mean of zone totals
$\hat{Y}_i$ = estimated total for zone i
N = number of zones in the population
n = number of zones in the sample
M_i = number of plots in zone i population
M = total number of plots in the population
m_i = number of sample plots within zone i
$\overline{y}_i$ = estimated mean for zone i
S_i^2 = estimated variance for zone i
Y_{ij} = observation from zone i,plot j
$\overline{y}$ = estimated mean of plot observation
f_1 = finite population correction factor for zone sample = $^n/_N$
f_2 = finite population correction factor plot = $\sum m_i/\sum M_i$
ps = post-stratified estimated
H = number of strata

large ($n_h \geq 20$, Cochran 1977); a condition that is satisfied within the parameters of this study. Also, it is worth noting that stratification of zones may help reduce variance during the initial SUUS implementation phase, but there is no advantage of doing so once all zones are measured as there is no longer a sample of the primary units.

Costs

Quantifying the inventory cost differences between SRS and SUUS panelized designs is difficult due to the many factors affecting operational implementation. In this analysis, we held some of the design factors constant (sampling intensity, plot size, plot configuration); however, other logistical factors such as SUUS zone size, the number, size, and location of field crews, modes of transportation and associated corridors, and per diem expenses for travel also have a large impact on the cost structure. With this in mind, we estimated the cost differences for field work based on approximate FIA field crew staffing levels, travel distances, and per diem rates in the study area. The cost models presented are general in nature such that they are applicable to other areas having different characteristics such as road density, wages, fuel costs, etc. Other cost pools such as administration, information management, and analysis and reporting were considered to be the same under either design.

For the cost analysis in our study area, we assumed each field crew was composed of one person whose home base was located near the center of a geographic area of responsibility (GAR). Size of GARs was determined by dividing the entire study area by the number of crews in the area. Based on current FIA staffing levels in this area, the number of crews was set at 12. To simplify analyses, the area for a crew member was represented by a circle with a radius of approximately 96.4 km (Fig. 2). We assumed that all distances to plots were linear from the point of origin. The number of plots that can be completed from the home base location depends on

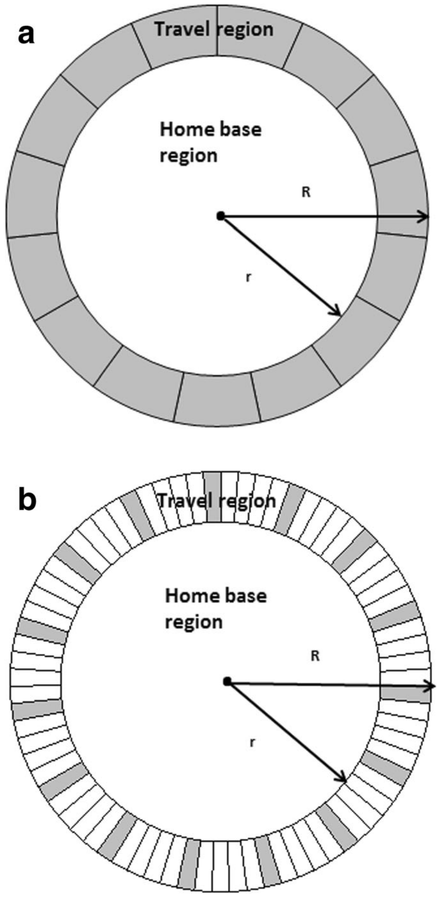

Fig. 2 A geographic area of responsibility (GAR) for one field crew. The dot depicts the home base location; *inner circle* represents the area where plots will be completed from home base; *shaded area* in travel region depicts the zones to be visited in a single panel that require overnight travel to complete plots. **a** SRS scenario where zone sizes are larger due to increased inter-plot distance, and each zone must be visited in a panel. **b** SUUS design where zones are smaller (closer inter-plot distances), and only a subset are visited within a given panel. The distances r and R are shown for the computation of average distance to zones in the travel region using Eq. (9)

various factors, including average day light length, practical field work logistical constraints, and administrative requirements. As such, a range of possible distances (10–90 km in 5-km increments) was evaluated within which the crew member would return home every night from field work. Field work outside the specified distance would require the crew person to stay overnight (Fig. 2). The distance threshold was used to determine travel costs for field work done in travel status as well as from home base. In both cases, the cost components were vehicle costs and crew salary.

Average travel distances were computed using the following equation to account for the circular areas of responsibility (notational definitions are given in Table 3):

$$\overline{\mathrm{D}}_{\bullet} = \frac{\int_r^R 2\pi r^2 dr}{\int_r^R 2\pi r dr} = \frac{\left(2/3\right)\pi R^3 - \left(2/3\right)\pi r^3}{\pi R^2 - \pi r^2} = \frac{2\left(R^3 - r^3\right)}{3\left(R^2 - r^2\right)} \tag{9}$$

For example, if the GAR has R=96.4 km and a distance threshold of r=80 km is specified, this results in $\overline{D}_t$ = 88.4 km. For the average travel from home base, R becomes 80 and r=0, so $\overline{D}_b$ = 53.3 km.

Total distance traveled by a crew is equal to the time spent traveling from home base to the zones beyond the threshold, the travel from the lodging to each plot, and the travel to plots near home base:

$$\mathrm{Dist} = T_t\left(2\overline{D}_t + 2\overline{I}\left(\overline{m} - 1\right)\right) + T_b\left(2\overline{D}_b\right) \tag{9.1}$$

The average area of each zone can be used to compute the average driving distance between the lodging and each plot within the zone:

$$\overline{I} = \frac{2}{3}\sqrt{\frac{\overline{A}}{100\pi}} \tag{9.2}$$

Due to the requirement that only plots in the same panel can be measured in a given year, the average distance between plots ($\overline{I}$) for SRS is the square root of the number of panels multiplied by the grid point separation distance, which in the case of a five-panel cycle and a 5-km grid is $5\sqrt{5} = 11.2$ km

Table 3 Notational definitions for Eqs. (9)–(13)

C_v = cost to operator vehicle ($)
C_t = salary cost while operating vehicle ($)
C_p = salary cost to measure plots ($)
C_d = daily lodging and per diem costs ($)
V_{rate} = vehicle cost per km ($)
T_t = number of overnight trips
$\overline{D}_t$ = average distance to travel zone (km)
$\overline{I}$ = average inter-plot distance within a zone
$\overline{m}$ = average number of plots per zone
$\overline{A}$ = mean number of hectares per zone
T_b = number of plots measured from home base
$\overline{D}_b$ = average distance to home base plot (km)
$\overline{V}$ = average travel speed (km/h)
S_c = crew salary cost ($/h)
F_p = proportion of area that is forested
t_f = time to measure forest plot
t_n = time to measure nonforest plot
$\overline{L}$ = average daily lodging rate
$\overline{P}$ = average daily per diem rate
R = outer GAR circle radius
r = inner GAR circle radius
Dist = total travel distance within GAR

Vehicle costs per crew can be generally described by

$$C_v = V_{\mathrm{rate}}\mathrm{Dist} \tag{10}$$

where the notational definitions for Eqs. (9–13) are given in Table 3.

For the SRS design, it was assumed that if a crew traveled outside the distance threshold, they would stay for an entire week and complete five plots. Thus, from a cost calculation standpoint, the SRS design was considered to have $\overline{m}$ = 5 for field work requiring travel. The SUUS design was evaluated for $\overline{m}$ = 5, 10, 15, and 20.

Crew salary costs for time spent traveling were calculated using

$$C_t = \frac{\mathrm{Dist}\ S_c}{\overline{V}} \tag{11}$$

Two components determine the total distance traveled. For the travel portion of fieldwork, there is the distance to the zone as well as travel between plots within the zone. Field work done from home base only includes the daily trips to plots.

Two cost pools were considered to be invariant to sample design: (1) the cost to measure individual plots and (2) lodging and per diem expenses. However, it was important to quantify these pools such that valid total cost estimates could be produced. For individual plot measurements, the costs were assessed in the context of crew salary multiplied by the number of plots and the average measurement time for forest and nonforest plots.

$$C_p = S_c\left(T_t\overline{m} + T_b\right)\left(F_p t_f + \left(1 - F_p\right)t_n\right) \tag{12}$$

For this study, it was assumed a nonforest plot would require 0.5 h, while a forested plot would take 5 h to complete (including walking from vehicle to plot and back to vehicle). The number of plots in each category was determined using an estimate of the proportion of the study area containing forest land and the total sample size. In this specific analysis, the proportion of forest land was 0.46 and the total number of plots was 14,619. This information was used to compute the total number of hours spent on plot and the associated cost based on crew salary per hour.

Because the same number of plots was completed in the travel region regardless of design, the actual time to complete the work was approximately the same, and thus crew lodging and per diem costs were assumed to be the same under all scenarios.

$$C_d = T_t\left(\overline{L} + \overline{P}\right)\overline{m} \tag{13}$$

For this study, the daily lodging cost was assumed to be \$85, the per diem amount was \$50, and these costs were incurred for each plot measured in the travel region of the GAR. To provide context to the cost Eqs. (9–13), an example is provided where the per crew costs shown are for the SUUS design having zones of average size $\overline{m}$ = 10 ($\bar{A} \approx 2430$ ha, $\bar{I} \approx 5.8$ km), a GAR of 96.4 km with travel distance threshold of 80 km, and an average travel speed of 40 km/h. Under this scenario, each crew would visit 38 zones in travel status and complete 85 zones from the home base. The total costs for a single crew over the five panels by component (vehicle, travel, plot measurement, and daily expense) are

$$\begin{aligned}
\text{Dist} &= T_t\left(2\overline{D}_t + 2\overline{I}\left(\overline{m} - 1\right)\right) + T_b\left(2\overline{D}_b\right) = 38[2(88.4) + 2(5.8)(10-1)] + 850[2(53.3)] = 99,312\text{km}\\
C_v &= V_{\text{rate}}\text{Dist} = 0.35\ (99,312) = \$34,759\\
C_t &= \frac{\text{Dist}\ S_c}{\overline{V}} = \frac{(99,312)\ 25}{40} = \$62,070\\
C_p &= S_c\left(\overline{m}(T_t + T_b)\right)\left(F_p t_f + \left(1 - F_p\right)t_n\right) = 25(10(38 + 85))(0.46(5) + (0.54)0.5) = \$79,028\\
C_d &= T_t\left(\overline{L} + \overline{P}\right)\overline{m} = 38\ (85 + 50)10 = \$51,300
\end{aligned}$$

Thus, total SUUS field work cost for 12 crews would be 12*(\$34,759+\$62,070+\$79,028+\$51,300)=\$2,777,184. As noted earlier, the costs for the SRS panelized design were computed under the assumption $\overline{m} = 5$; however, under the SRS design, each zone essentially needs to be visited every year because each zone will have plots belonging to all five panels.

Results and discussion

Estimation

A primary point of interest regarding the SUUS design was the loss of precision relative to SRS in the years prior to full implementation when all zones are measured. When compared to SRS estimates, the SUUS standard errors when only a single panel (20 %) has been sampled were approximately 1.5–2.2 and 1.4–2.1 times larger for forest area and biomass, respectively (Table 4). A relevant characteristic of the SUUS design is that larger standard errors are associated with larger zone sizes. Larger zone sizes result in a smaller number of PSUs, which make the estimates less precise. However, large zones also tend to have smaller within-zone variance in comparison to smaller zones due to larger within-zone sample size. These results indicate that reduced within-zone variance does not compensate for the reduced zone sample size, and thus, the precision of estimates will suffer as zone size increases.

As the number of sampled panels increases over the years, the relative precision difference decreases. For example, with four of five panels included, SUUS standard errors for forest area and biomass were 1.05–1.34 and 1.03–1.27 times larger, respectively (Table 4). When all five panels were included, the SUUS standard error was slightly smaller than the corresponding SRS standard error, which is attributed to the zones acting as strata at full implementation and the geographic grouping of plots providing a small variance reduction due to spatial autocorrelation of plots within zones. Thus, the option of estimation under an SRS paradigm when all zones have been measured is generally not preferred in comparison to retaining the SUUS estimator. A general property of stratified estimation is the rate of reduction in variance decreases rapidly as the number of strata increases, so having numerous strata may be excessive and loss of degrees of freedom may be a concern in some cases; however, there is no technical reason strata defined by zones could not be employed.

Zones acting as strata at full inventory implementation are essentially a byproduct of measuring all zones in the population. This differs from traditional stratified designs where strata are deliberately chosen to increase the precision of estimates. In this study, zones were grouped into canopy cover strata to assess how much further variance reduction might be achieved. When only one panel had been measured, the reduction in sampling error ranged from 31 to 48 % for estimates of forest land area (not shown) and 23 to 41 % for estimates of biomass (Fig. 3). As expected, increasing the number of strata improved the precision of the estimates; however, at a decreasing rate. The effectiveness of the stratification also depended on the number of measured panels; as more panels were measured, the variance reduction decreased (Fig. 4). This is largely due to stratification only improving the first term of (2), which has a decreasing contribution as the proportion of zones that has been measured increases. This result suggests that the relative difference in sampling errors between SUUS and SRS can be decreased for the first few panels by employing stratification of zones.

The plot stratification applied to SRS consistently resulted in decreases of 28–30 % in standard errors when compared to non-stratified SRS estimates. For estimates of both forest land area and biomass, the post-stratified SRS design consistently provided smaller standard errors than the SUUS design with stratification of zones. The smallest relative differences between the two designs occurred when all panels were measured, where standard errors were approximately 20 % smaller for forest land area and 10 % smaller for biomass. Thus, from a precision standpoint, it is better to stratify plots under the SRS design than to stratify zones under the SUUS framework. Similarly, when no stratification is employed, SRS is also preferable with the exception of when all panels are measured and the SUUS estimates are more precise due to zones acting as strata.

An alternative stratification scheme could be employed where plots within zones are stratified. For the zone sizes considered here, this is problematic as the small sample sizes can produce biased estimates of precision (Westfall et al. 2011). This is a conundrum because smaller zone sizes have relatively large within-zone variance, and thus, these small zones would likely benefit the most from plot stratification. As zone size increases, the value of stratifying plots decreases. In an analysis (not shown) that used much larger zones that could support adequate stratum sample sizes, there was little advantage to stratifying plots within zones prior to full implementation where all zones are measured. This was due to the variance being largely driven by the first term in (2), which accounts for variation among the zones. The analysis showed that this term usually contributed 90 % or more of the total variance. The stratification of plots within zones only affected within-zone variance, which was small relative to between-zone variance.

Another consideration for the SUUS design is estimation of current values versus trends. As discussed earlier, when the inventory cycle is complete and all plots have been measured, there is an opportunity to abandon SUUS from an estimation standpoint and treat the design as SRS or post-stratified random sampling. However, this only applies to current estimates. For trend estimates, the SUUS design will still apply until all the plots have been remeasured, e.g., change observations will occur for panel 1 only when panel 1 is remeasured and the zones are still the PSU for observations of change between inventories. As such, the use of SUUS estimators must continue for at least two full inventory cycles. This may influence design choice, as trend estimates tend to be highly variable even under a stratified SRS design, and the precision using SUUS will likely be poor initially but will improve with each remeasurement panel until the end of the second cycle.

Table 4 Sample size and sampling error percent for estimates of total forest area and biomass by number of panels in the sample for SRS and SUUS designs with various zone sizes

No. of panels	SRS			5/zone			10/zone			15/zone			20/zone		
	n	Area SE%	Biomass SE%	*n*	Area SE%	Biomass SE%	*n*	Area SE%	Biomass SE%	*n*	Area SE%	Biomass SE%	*n*	Area SE%	Biomass SE%
1	2929	1.87	2.50	587	2.84	3.49	296	3.26	4.05	195	3.81	4.64	145	4.21	5.13
2	5866	1.33	1.78	1174	1.83	2.28	591	2.08	2.60	390	2.40	2.96	291	2.64	3.23
3	8791	1.09	1.46	1762	1.33	1.70	885	1.49	1.89	585	1.69	2.11	436	1.84	2.28
4	11,707	0.95	1.27	2349	1.00	1.31	1180	1.08	1.41	779	1.19	1.52	582	1.27	1.61
5	14,619	0.85	1.14	2937	0.73	1.01	1476	0.73	1.01	973	0.73	1.02	728	0.74	1.02

Costs

Some of the key factors influencing cost are the travel speed, salary rate, zone size, and travel distance threshold. The relationships among these factors can be investigated to provide further insight into cost drivers. For example, the combined effects of salary rate and travel speed indicate that salary rate has a substantial impact on travel costs ($C_v+C_t+C_d$) when travel speeds are low; however, salary rate becomes considerably less important as travel rate increases (Fig. 5). Similarly, low travel speeds are most costly when the zone size is small because more round trips are needed than for large-size zones (Fig. 6). It was also noted that the travel costs tend to flatten out as zone size increases, indicating that zone sizes larger than 20 may have minimal impact on reducing total travel cost. Other relationships can be generally inferred; however, quantification is recommended for any specific points of interest, e.g., how might low salary rates and high travel speed influence the costs at various travel distance thresholds.

In comparison to the SRS design, SUUS may be implemented at a lower cost because each zone is visited only once per inventory cycle (compared to annual visits under SRS). In addition to the travel costs associated with zones beyond the distance threshold, the total field cost includes the travel to home base plots and time spent measuring all plots in the inventory. As noted earlier, the SRS design would likely be implemented with a zone size of 5 plots for areas located in the travel region, and thus, it may be expected that costs would be the same for both SRS and SUUS at this zone size. However, the inter-plot distance between plots in the zone under SRS is the square root of 5 times larger. At the 80-km distance threshold, assuming travel speed is 40 km/h and salary rate is \$25/h, the cost to implement SRS is approximately 2.5 % higher than the SUUS design.

Of particular interest are the overall strategies for minimizing cost for the entire field work effort via implementation of the SUUS design. This primarily entails specifying optimal zone size and travel distance threshold parameters. Often other factors such as salary costs and travel speeds are beyond the control of inventory planners. However, these factors are important when determining the design parameters. An examination of total field costs in the context of travel distance thresholds and travel speeds shows that travel speed should be a consideration in determining the distance threshold. Figure 7 indicates that the lowest cost

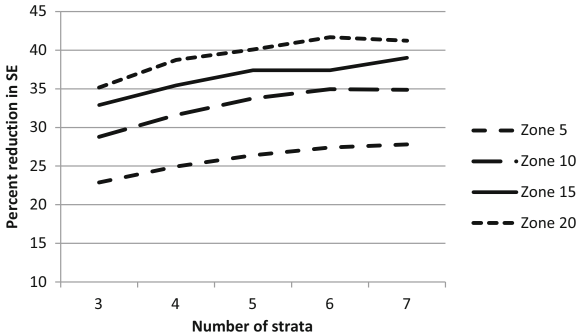

Fig. 3 Percent sampling error reduction due to stratification of zones (vs. unstratified zones) for estimates of biomass by zone size and number of strata

(assuming zone size=5, salary=$25/h) when travel speeds are near 10 km/h occurs at a distance threshold of 30 km, whereas the cost minimum occurs at 75 km for a travel speed of 30 km/h. When 50 km/h is attained, the results suggest that no zones should be used at all for the GAR analyzed in this study. Essentially, the efficiency afforded by traveling at that speed makes it more economical to spend time traveling back to home base than to incur per diem costs to stay overnight.

A similar pattern is found when evaluating distance thresholds under various salary costs. Assuming a basis of zone size=5 and travel speed=40 km/h, high salary costs ($60/h) suggest a distance threshold of approximately 50 km provides the most efficient solution (Fig. 8). As salary costs are reduced to $40/h, a distance threshold of 65 km is near optimal. Similar to the travel speed analysis, when travel becomes relatively inexpensive due to low salary costs (e.g., $20/h), it becomes more cost-effective to spend time driving throughout the GAR than allowing per diem expenses.

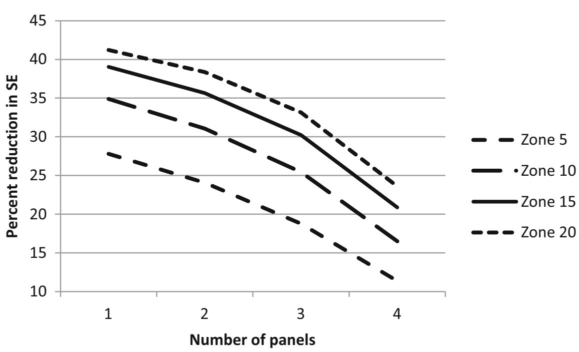

Fig. 4 Percent sampling error reduction due to stratification of zones (vs. unstratified zones) for estimates of biomass by zone size and number of panels measured when zones are grouped into seven strata

The overall percentage of savings associated with employing the SUUS design ultimately depends on numerous factors. To provide some context to the findings, the percentage of savings for an inventory cycle was evaluated over the 4 zone sizes and range of distance thresholds, assuming the travel speed=40 km/h and salary=$25/h. The largest savings were realized at small distance thresholds as considerable travel to zones is incurred; travel costs to and from zones was the cost component that SUUS primarily influenced (Fig. 9). It also showed that increasing zone size results in further cost reduction. Increasing zone size also resulted in higher sampling errors (Table 4), exemplifying the traditional cost vs. precision conundrum. A notable outcome was the relatively large difference in savings between zone sizes of 5 and 10 plots, and the subsequent relatively small differences as zone size increased beyond 10 plots. Thus, a reasonable compromise between cost and precision may be most appealing at zone size 10; however, planners need to further evaluate this suggestion in the context of their specific inventory cost components and population structure. At distance thresholds likely to be employed (i.e., those that result in most plots being done from home base), a 2–7 % decrease in field costs may be realized. The actual percentage reduction for the entire inventory will be less, depending on the organizational formation and cost components for various other tasks such as programming, data processing, pre-field work, reporting, and administrative support.

Clearly, this analysis represents only one of nearly limitless forest inventory situations that may be encountered, and simulating any given scenario necessarily requires some practical assumptions. To better understand how SUUS may be advantageous, planners are encouraged to undertake analyses pertinent to their circumstances. The cost of traveling to zones is the key determinant of savings that SUUS might confer. This cost is comprised of a combination of salary and vehicle expenses. Generally, the vehicle expense is the primary driver of realized savings when wage costs are low; however, the influence of vehicle expense diminishes as wage costs increase. The largest savings are realized when wage costs are low and travel times are long and relatively expensive per unit of distance, a situation most likely to be encountered in areas with limited transportation and infrastructure networks, and poor accessibility

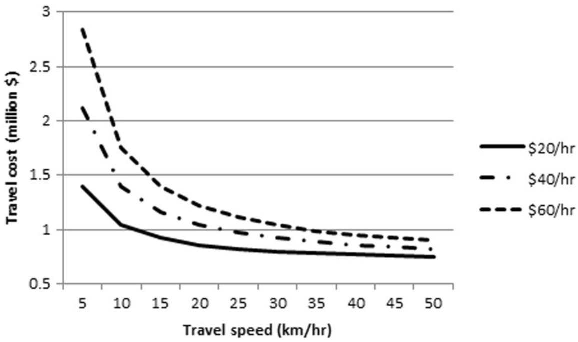

Fig. 5 Field work travel cost 12* (C_v+C_t+C_d) for travel region by salary cost and travel speed for distance threshold=80 km and zone size=5

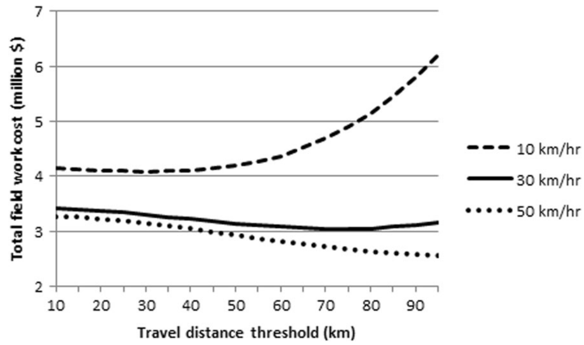

Fig. 7 Total field work cost (C_v+C_t+C_d+C_p) for various travel speeds and distance thresholds for zone size=5 and salary=$25

such as is common in many developing countries. In these types of situations, the transportation costs can be much higher. For example, boats are often needed because rivers provide the only access routes, and considerable walking time may be necessary between the river shore and plot location. Further, the remoteness and type of terrain often requires a larger number of crew members and additional equipment and supplies, and thus, the cost per unit of distance traveled increases substantially. As transportation costs increase and roundtrip times increase, substantial reductions in field work costs can be realized by visiting the area only once per cycle. More complex evaluations may be needed to evaluate trade-offs between SUUS and existing designs or when developing new SUUS-based inventories where optimal cost/precision determinations are required. We refer readers to Hansen et al. (1953) for a comprehensive overview of design considerations and cost accounting options.

Conclusion

During the initial implementation of the SUUS design, the typical compromise between cost and precision is evident, i.e., reducing cost equates to less precision. Thus, prior to the completion of the inventory cycle, the precision of estimates suffers in comparison to the SRS design. However, at the end of the first measurement cycle for current estimates and at the end of the second cycle for trend estimates, there is the option to employ SRS estimation methods. Under a scenario where no stratification is used, it is better to continue using SUUS estimators as each individual zone acts as a "pseudo-stratum" due to spatial correlation of plots within zones and more precision is attained than with the unstratified SRS design. If the SUUS design is chosen and additional information is available for stratification purposes, stratifying zones into relatively homogenous categories and using associated stratified estimation methods is

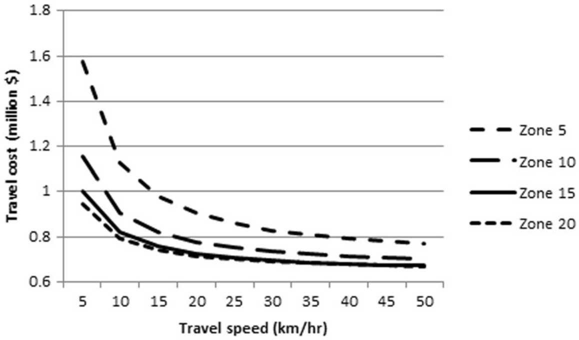

Fig. 6 Field work travel cost 12* (C_v+C_t+C_d) for travel region by zone size and travel speed for distance threshold=80 km and salary=$25/h

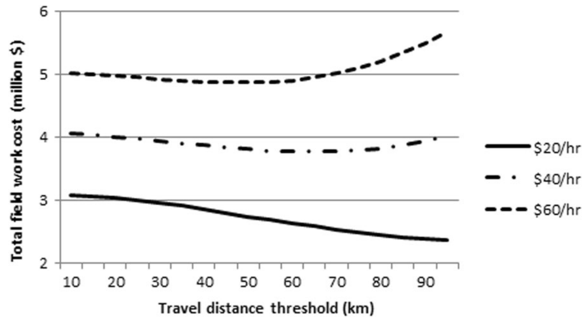

Fig. 8 Total field work cost (C_v+C_t+C_d+C_p) for various salary rates and distance thresholds for zone size=5 and travel speed=40 km/h

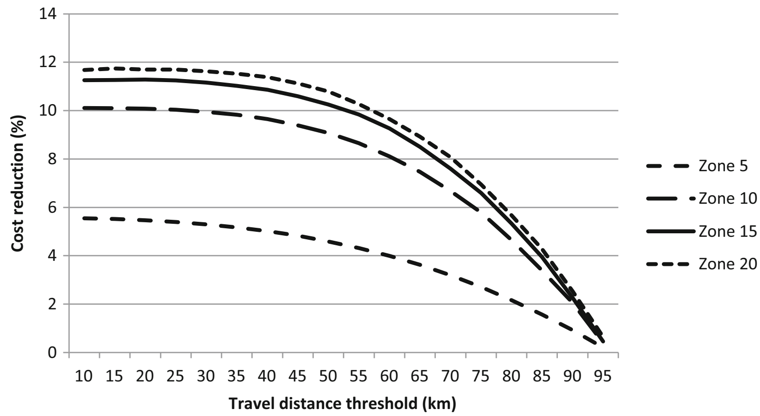

Fig. 9 Percent reduction in total field work cost by zone size and distance threshold for travel speed=40 km/h and salary=$25/h

recommended until all panels are measured. Stratified SRS estimation should be implemented when the inventory cycle is completed in order to enhance precision over stratified SUUS. This suggests an appropriate long-term strategy to optimize precision at a reduced cost is to continually employ the SUUS design to minimize field work costs while adopting stratified SRS estimation to maximize precision.

Cost efficiency depends on numerous factors; this study primarily focused on travel distance thresholds and zone sizes as design parameters. Optimal specification of these parameters in terms of cost reduction was primarily driven by travel speeds and salary costs. Of particular note was the existence of a travel efficiency level (low salary, high travel speed) that resulted in lower costs for the SRS design than the SUUS design (Figs. 6 and 7). However, such efficiencies are unlikely to be realized in practice, and thus, the SUUS design will generally be preferred economically. For distance thresholds thought to be logistically practical, field work cost savings in the range of 2–7 % may be realized by using the SUUS design. The cost savings with respect to zone size suggested greater cost reductions were associated with larger zone sizes; however, the relative differences decreased rapidly as zone size increased and a zone size of 10 was considered a balanced compromise between cost and precision for this study.

Planners and analysts need to judge the practical and political implications of employing the SUUS design when establishing a new forest inventory. Selection of a design that has relatively poor precision during the initial years may be difficult for consumers of the data to accept. Available funding and interest in the results within the host organization (and often the broader forestry community) is often highest during the first few years of inventory implementation and also in the initial years of remeasurement when estimates of trends become available. Consideration of both current funding levels and the potential for long-term funding should be a factor in design selection. Given the goal of performing estimation under a stratified SRS paradigm once all the zones have been measured (for both current and change estimates), it would be problematic if full measurement was not attained, as only the SUUS estimates having lower precision would be available. Also, SUUS may be more difficult to implement, e.g., field support, planning, data management, and analysis functions would be more complex than the SRS design. Full consideration of the effects on all aspects of the inventory program should be undertaken prior to implementation of the SUUS design.

Acknowledgments The authors are thankful to Vicente Monleon, John Stanovick, Christoph Kleinn, and several anonymous reviewers for providing valuable comments to improve the manuscript.

References

Bechtold, W.A., & Patterson, P.L. (Eds.). (2005). The enhanced forest inventory and analysis program-national sampling design and estimation procedures. U.S. For. Serv. Gen. Tech. Rep. SRS-80.

Bechtold, W.A., & Scott, C.T. (2005). The forest inventory and analysis plot design. In W.A. Bechtold & P.L. Patterson (Eds.), The enhanced forest inventory and analysis program–national sampling design and estimation procedures. U.S. For. Serv. Gen. Tech. Rep. SRS-80. pp. 27–42.

Brassel, P., & Lischke, H. (Eds.). (2001). *Swiss national forest inventory: methods and models of the second assessment*. Birmensdorf: WSL Swiss Fed. Res. Inst.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: Wiley.

Gillis, M. D., Omule, A. Y., & Brierley, T. (2005). Monitoring Canada's forests: the national forest inventory. *Forestry Chronicle, 81*(2), 214–221.

Gregoire, T. G., & Valentine, H. T. (2008). *Sampling strategies for natural resources and the environment*. Boca Raton: Chapman & Hall/CRC Applied Environmental Statistics.

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sample survey methods and theory*. New York: Volume I. John Wiley and Sons, Inc.

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering and Remote Sensing, 70*(7), 829–840.

Kändler, G. (2009).The design of the second German national forest inventory. In R.E. McRoberts, G.A. Reams, P.C. Van Deusen, & W.H. McWilliams (Eds), Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium U.S. For. Serv. Gen. Tech. Rep. WO-79. pp. 19–24.

Kangas, A., & Maltamo, M. (Eds.). (2006). *Forest inventory: methodology and applications*. Dordrecht: Springer.

Köhl, M., & Scott, C.T. (1992). Survey planning for national forest inventories. In G.B Wood & G.J. Turner (Eds.), Proceedings of the Integrating Forest Information over Space and Time Conference ANUTECH Pty Ltd. pp. 325–329.

Köhl, M., Lister, A., Scott, C. T., Baldauf, T., & Plugge, D. (2011). Implications of sampling design and sample size for national carbon accounting systems. *Carbon Balance and Management, 6*, 10. doi:10.1186/1750-0680-6-10.

Lister, A. J., & Scott, C. T. (2009). Use of space-filling curves to select sample locations in natural resource monitoring studies. *Environmental Monitoring and Assessment, 149*, 71–80.

Loetsch, F. (1957). A forest inventory in Thailand. Unasylva 1(4). Available from http://www.fao.org/docrep/x5385e/x5385e06.htm. Accessed 10 July 2015.

McRoberts, R. E., Tomppo, E. O., & Næsset, E. (2010). Advances and emerging issues in national forest inventories. *Scandinavian Journal of Forest Research, 25*(4), 368–381.

Reams, G.A., Smith, W.D., Hansen, M.H., Bechtold, W.A., Roesch, F.A., & Moisen, G.G. (2005). The forest inventory and analysis sampling frame. In W.A. Bechtold & P.L. Patterson (Eds.), The enhanced forest inventory and analysis program—national sampling design and estimation procedures U.S. For. Serv. Gen. Tech. Rep. SRS-80. pp. 11–26.

Schreuder, H.T., Ernst, R., & Ramirez-Maldonado, H. (2004). Statistical techniques for sampling and monitoring natural resources. U.S. For. Serv. Gen. Tech. Rep. RMRS-GTR-126.

Stehman, S. V., Wickham, J. D., Fattorini, L., Wade, T. D., Baffetta, F., & Smith, J. H. (2009). Estimating accuracy of land-cover composition from two-stage cluster sampling. *Remote Sensing of Environment, 113*(6), 1236–1249.

Tokola, T., & Shrestha, S. M. (1999). Comparison of cluster-sampling techniques for forest inventory in southern Nepal. *Forest Ecology and Management, 116*, 219–231.

Tomppo, E., Gschwantner, T., Lawrence, M., & McRoberts, R. E. (Eds.). (2010). *National forest inventories-pathways for common reporting*. New York: Springer.

Westfall, J. A., Patterson, P. L., & Coulston, J. W. (2011). Post-stratified estimation: within-strata and total sample size recommendations. *Canadian Journal of Forest Research, 41*(5), 1130–1139.