European Journal of Forest Research (2019) 138:375–382
https://doi.org/10.1007/s10342-019-01171-9

ORIGINAL PAPER

Double sampling for post-stratification in forest inventory

James A. Westfall[1] · Andrew J. Lister[1] · Charles T. Scott[2] · Thomas A. Weber[1]

Received: 19 July 2018 / Revised: 20 December 2018 / Accepted: 24 January 2019 / Published online: 14 March 2019

Abstract

Many national forest inventories (NFI) use auxiliary data to increase the precision of estimates. Typically, this is accomplished via stratified estimation techniques that rely on assignment of similar sample plot observations to strata constructed with the goal of lowering the variance of estimates. While early applications of stratification used strata constructed from photo-interpretation of aerial photography, current technology makes using wall-to-wall digital map information more appealing due to automated processing capabilities; however, there is generally a reduction in classification accuracy in comparison with photo-interpretation and a concomitant decrease in the precision of estimates. While most established NFI have permanent plots and employ post-stratification (PS) with stratum weights known from a map, it is unclear what are the compromises compared to using a photo-interpretation (PI) approach. In this study, differences in cost and precision were evaluated for post-stratification using strata derived from a digital map and double sampling for post-stratification (DSPS) with strata created from PI of aerial imagery. It was found that DSPS consistently provided better precision than PS for estimates of total biomass and forestland area with approximately 13 PI points per sample plot, which incurred a cost increase equivalent to 0.5% per ground plot. Increasing the number of PI points per plot resulted in further gains in precision, with cost increases proportional to the PI intensity. To attain specific precision goals, DSPS was generally less costly than increasing the sample size under PS, although the PS design was more cost-effective if the PI intensity was too low. The results of this study provide a decision framework for inventory planners considering sampling designs that rely on post-stratified estimation.

Keywords Photo-interpretation · Estimator variance · Cost:precision · Remote sensing

Introduction

Stratification in sampling design and subsequent estimation to achieve increased precision of estimated population parameters is well known to forest inventory practitioners (Frayer and Furnival 1999; McRoberts et al. 2002; Kangas and Maltamo 2006). Although a priori specification of the strata is often preferred, it is often ill-advised for ongoing large-area inventory and monitoring programs having permanent plots such as national forest inventories (NFI), due to the likely increase in within-stratum heterogeneity over time and the associated loss of precision (De Gruijter et al. 2006). It is arguably more desirable to employ post-stratification where the strata reflect changing landscape conditions in subsequent inventories, even though there is a slight increase in estimator variance from doing so (Cochran 1977). Currently, it is common to use wall-to-wall satellite imagery for defining strata used in forest inventory estimation (Tomppo et al. 2008). This process is efficient because it can be largely automated and the wall-to-wall map coverage affords known stratum weights. Before the advent of digital mapping with satellite imagery, double sampling for stratification was often used via photo-interpretation (PI) of selected points (Bickford 1952; Lam et al. 2011). Disadvantages to this approach were cost (time consuming PI work) and the use of double sampling introduced additional variance arising from stratum weights being estimated instead of known values. However, an advantage of using PI over digital mapping was increased accuracy of stratum assignments due to the use of a human interpreter, which can lead to more precise estimates than post-stratification based on

Communicated by Arne Nothdurft.

✉ James A. Westfall
jameswestfall@fs.fed.us

1 U.S. Forest Service, Northern Research Station, Newtown Square, PA 19073, USA

2 West Chester, PA 19380, USA

automated pixel classifications (Hansen and Wendt 2000; Wayman et al. 2001).

Generally, the goal of optimal forest inventory design is to obtain estimates that meet specified precision goals for the least cost (Köhl et al. 2011). Thus, effective and efficient inventories are often determined via comparison of various cost pools and precision outcomes for candidate sampling designs (Westfall et al. 2016). The two cost pools directly related to precision of the estimates pertain to the number and cost of sample plots and the development of a stratification scheme for estimation purposes. Greater precision is obtained by increasing the sample size of the field plots, with the cost of the increased precision directly related to the cost of additional field work. A related cost:precision issue is the expense of developing the stratification information and the amount of precision gained relative to that cost. If the locations of field plots are determined prior to the stratification, stratified designs would rely on post-stratification. In these situations, comparisons of typical wall-to-wall post-stratification (PS) to the double sampling for stratification paradigm would therefore require the concept of double sampling for post-stratification (DSPS) to be considered. Specifically, the objectives of this study include: (1) derivation of the DSPS variance estimator, (2) evaluation of the cost:precision compromises for different levels of PI effort and (3) comparisons of the cost and precision between the PS and DSPS approaches.

Methods

Estimators

Parameters to be estimated are the population total $\hat{T}$ and its variance $\hat{V}(\hat{T})$. Regardless of the stratification paradigm being employed, the estimate of $\hat{T}$ is given by

$$\hat{T}_{\mathrm{PS}} = \hat{T}_{\mathrm{DSPS}} = A \sum_{\mathrm{h}=1}^{H} w_{\mathrm{h}} \bar{y}_{\mathrm{h}} = A\bar{y} \tag{1}$$

where A total area of the population (ha) w_{h} the estimated or known weight of stratum h, $\bar{y}_{\mathrm{h}}$ the sample mean of observations in stratum h ($h = 1\ldots H$ denote the strata) and $\bar{y}$ overall sample mean.

The variance of the total for post-stratification, $\hat{V}(\hat{T}_{\mathrm{PS}})$, has been derived in various sampling texts and is used in the FIA program, without correction for a finite population, as (Scott et al. 2005):

$$\hat{V}(\hat{T}_{\mathrm{PS}}) = A^2 \left[\sum_{\mathrm{h}=1}^{H} w_{\mathrm{h}} \frac{s_{\mathrm{h}}^2}{n} + \sum_{\mathrm{h}=1}^{H} \left(1 - w_{\mathrm{h}}\right) \frac{s_{\mathrm{h}}^2}{n^2} \right] \tag{2}$$

where n total number of plots and s_{h}^2 sample variance for stratum h.

To derive the estimator for $\hat{V}(\hat{T}_{\mathrm{DSPS}})$, the adaptation of equation 12.32 from Cochran (1977) as presented by Scott et al. (2005) for double sampling for stratification (DSS) serves as the basis:

$$\hat{V}(\hat{T}_{\mathrm{DSS}}) = A^2 \left[\sum_{\mathrm{h}=1}^{H} \left(\frac{n'_{\mathrm{h}} - 1}{n' - 1} \right) \frac{n'_h}{n'} \frac{s_h^2}{n_{\mathrm{h}}} + \frac{1}{n' - 1} \sum_{\mathrm{h}=1}^{H} \frac{n'_{\mathrm{h}}}{n'} \left(\bar{y}_{\mathrm{h}} - \bar{y} \right)^2 \right] \tag{3}$$

The previously undefined notation includes the following: n' number of first phase PI points, n'_{h} number of first phase PI points in stratum h and n_{h} number of plots in stratum h. The first term estimates the stratified variance under proportional allocation, while the second term is due to estimation of the stratum weights. In a post-stratification context, additional variance is incurred due to the variability of within-stratum sample sizes, n_{h}. Specifically, the estimator in (3) needs to incorporate the expected value of the number of observations per stratum (Stephan 1945):

$$E\left(\frac{1}{n_{\mathrm{h}}} \right) = \frac{1}{n w_{\mathrm{h}}} + \frac{1 - w_{\mathrm{h}}}{n^2 w_{\mathrm{h}}^2} \tag{4}$$

Substituting the right-hand side of [4] for $\frac{1}{n_{\mathrm{h}}}$ in (3) suggests the form of the estimator for $\hat{V}(\hat{T}_{\mathrm{DSPS}})$:

$$\begin{aligned} \hat{V}(\hat{T}_{\mathrm{DSPS}}) = A^2 \Bigg[& \sum_{\mathrm{h}=1}^{H} \left(\frac{n'_{\mathrm{h}} - 1}{n' - 1} \right) \frac{s_{\mathrm{h}}^2}{n} + \sum_{\mathrm{h}=1}^{H} \left(\frac{n'_h - 1}{n' - 1} \right) \\ & \times \frac{\left(1 - w_{\mathrm{h}}\right) s_{\mathrm{h}}^2}{n^2 w_{\mathrm{h}}} + \frac{1}{n' - 1} \sum_{\mathrm{h}=1}^{H} \frac{n'_{\mathrm{h}}}{n'} \left(\bar{y}_{\mathrm{h}} - \bar{y} \right)^2 \Bigg] \end{aligned} \tag{5}$$

The percent sampling errors of the estimates are calculated as:

$$\mathrm{SE}\%(\hat{T}_{.}) = \frac{\sqrt{\hat{V}(\hat{T}_{.})}}{\hat{T}_{.}} \times 100 \tag{6}$$

where $\hat{T}_{.}$ refers to either $\hat{T}_{\mathrm{PS}}$ or $\hat{T}_{\mathrm{DSPS}}$ depending on the design being analyzed.

In the context of post-stratification, it should be noted that $\hat{V}(\hat{T}_{\mathrm{DSPS}})$ as shown in (5) is the unconditional variance estimator, i.e., the variation across all combinations of within-strata sample allocations. There seems to be a longstanding debate among practitioners regarding whether the unconditional or conditional form of the post-stratified variance estimator is most appropriate for inference. Holt and Smith (1979) provide a brief synopsis of different presentations in the literature at that time and conclude from their research the conditional estimator is preferred. Other research by Durbin (1969), Valliant (1993) and Gregoire et al. (2016)

(among others) also suggests the use of the conditional formulation. Yet, some authors present the unconditional estimator as the basis for post-stratified inference (Scott et al. 2005; Köhl et al. 2006; McRoberts et al. 2013). The primary issue is that unconditional variance estimates will only be accurate when $W_h \approx n_h/n$ within strata. Generally, this should not be a concern for most NFI that implement a systematic or quasi-systematic sample design across the population, as sample sizes tend to be proportional to the size of the strata. For example, Saborowski and Cancino (2007) showed in a simulation of forest inventory using a post-stratified systematic design that nearly identical results are obtained from the conditional and unconditional variance estimators. Nonetheless, forest inventory specialists should be aware of circumstances that may favor a conditional approach.

Monte Carlo estimation of DSPS variance

To provide empirical evaluation of the variance estimator (5), a simulation study was conducted wherein a population containing 250,000 elements was generated. For each element, total tree biomass (kg/ha) and crown cover percent values were assigned. For biomass, the values were randomly chosen from a $N(\mu, \sigma)$ distribution where $\mu = 9200$ and $\sigma = 5500$ kg/ha. Similarly, the crown class percent values were chosen from a $U(0,1) \times 100\%$ distribution and rounded to the nearest 1%. To mimic a typical landscape, biomass values less than zero were assumed to represent nonforest land without trees and thus both biomass and crown cover were set to zero. It was also assumed that relatively large biomass values would occur in mature stands with crown closure, such that crown cover was assumed to be 100% when biomass exceeded 11,000 kg/ha. To implement stratified estimation, each population element was assigned to one of the three strata based on the crown cover: 1 = 0–9%, 2 = 10–50%, and 3 = 51–100%.

The Monte Carlo simulation was conducted by randomly selecting 100 sample points to represent field plot observations. Subsequently, an additional 900 sample points were chosen and the stratum weights were estimated from the set of 1000 points (w_h; double sampling for stratification). The 100 initial sample points were considered to have been established prior to the development of strata (post-stratification). At each iteration, the population total was estimated from (1) and the estimated variance calculated using (5). After 25,000 iterations, the mean value of $\hat{V}\left(\hat{T}_{\mathrm{DSPS}}\right)$ and the variance of the estimated totals $\hat{T}_{\mathrm{DSPS}}$ were calculated. The 0.1% difference between these two variance estimates suggests the specification of (5) provides a suitable estimator of the variance when a DSPS design is employed. Evaluation of the individual terms of (5) indicated proportions of the total variance were 0.949 for the first term, 0.013 for the second term, and 0.038 for the third term bracketed on the right-hand side of (5), respectively. Thus, for our example population, approximately 95% of the variance arises from plot-to-plot variability, whereas contributions due to random post-strata sample sizes (1.3%) and estimated stratum weights (3.8%) comprise the remaining 5%. The relative contributions of each variance component may differ for other population/stratification scenarios; however, it is expected the plot-to-plot variance will still be the primary source of uncertainty (Brown and Westfall 2012).

Application to NFI

To provide a case study of the implementation of the DSPS estimator and compare outcomes with those obtained from digital map post-stratification, data from the forest inventory and analysis (FIA) program within the US forest service (USFS) were used. Three counties that represent a range of areal extent and proportion of forestland area were chosen in eastern Pennsylvania (Fig. 1). The full cycle of inventory plots were measured over the period 2011–2016, under a systematic-unaligned sampling design having a sampling intensity of approximately 1 plot per 2428 ha (Reams et al. 2005). The field data were collected using a 0.067 ha 4-point plot design (Bechtold and Scott 2005) and include the proportion of plot area that is forested and individual-tree measurements used to predict tree biomass from models (U.S. Forest Service 2012). Plot-level observations of tree biomass are obtained by summing over all trees and expanding to a per-hectare basis. Although the field data were collected in a panelized design, this analysis combines all plots in the inventory cycle for estimation.

To develop the stratification scheme for PS, the national land cover dataset (NLCD) tree canopy map (Homer et al. 2015), a 30 m resolution, Landsat-based map depicting canopy cover proportion, was used to determine the strata and their associated weights. The pixel-based area for each canopy cover percent was determined from the NLCD map. Using these data and a minimum sample size rule of 10 plots per stratum (Westfall et al. 2011), canopy cover classes 0–5%, 6–50%, 51–75%, and 76–100% were chosen to define the post-strata. Due to the smaller sample size for Lehigh County, only three post-strata were created (0–5%, 6–75%, and 76–100%). The stratum weights were the canopy cover class area proportion of the total pixel-based area.

Estimated stratum weights for the DSPS design were obtained using national agriculture imagery program (NAIP) digital photography from 2015 (USDA-FSA-APFO 2016). Based on protocols presented by Bickford (1952), the PI point sample intensity was chosen to be approximately 25 points per FIA plot. PI points were distributed in a spatially balanced manner using a space-filling curve fractal. This method tessellated the region into the appropriate number of compact, equal-area subregions, within which one PI point

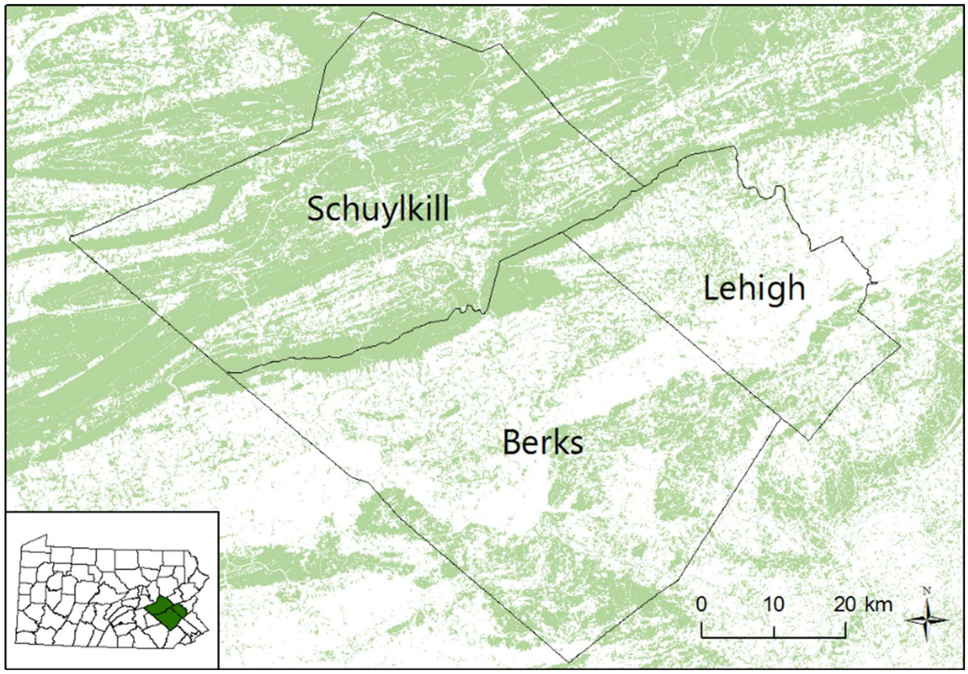

Fig. 1 Map of Berks (229,650 ha; 35% forested), Lehigh (90,987 ha; 31% forested) and Schuylkill (194,426 ha; 65% forested) Counties in eastern Pennsylvania, USA

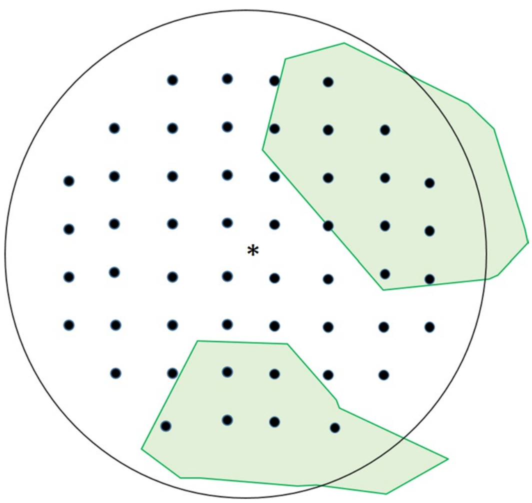

Fig. 2 Configuration of the PI point where crown cover (shaded areas) was assessed at 52 locations systematically arranged around the center (*) of a plot having 43.9 m radius. In this example, the crown cover is 18/52 = 34.6%

was randomly chosen (Lister and Scott 2009). The PI point that was closest to each FIA plot was removed from the PI analysis and replaced with the corresponding FIA plot location, which was also classified using PI. At each PI point, a circle with radius of 43.9 m was established (corresponding to the size of the outer boundary of an FIA plot; Fig. 2). Each circle contained a grid of 52 points to be assessed for intersections with tree crowns occurring in areas that would meet the FIA definition of forestland (U.S. Forest Service 2012). The canopy cover percent for the PI point is the percentage of the 52 grid points meeting the criteria (tree crown in forestland). The DSPS stratum weights were calculated by determining the proportion of PI points falling in each of the canopy cover classes (de Vries 1986). To make valid comparisons of results, the same canopy cover classes were used for stratification in both the PS and DSPS designs.

For both designs, estimates of total population biomass and forestland area in each county were calculated using the appropriate formulae described earlier. To better understand how the number of PI points affected the results, the amount of PI points was reduced by 25%, 50% and 75% and the estimates recalculated with each reduced PI sample. To provide an indication of expected effects of using fewer PI points, these reductions were replicated 1000 times and the average values were reported. The same process was repeated for reductions of 80%, 84%, 88%, 92% and 96% to examine the behavior of the estimates for relatively small PI efforts. Finally, a scenario was considered where the only PI information obtained was at the plot locations.

The estimation of costs for the DSPS design only included the variable costs associated with staff time to develop canopy cover information. A time stamp was generated in the data as each PI point was completed. Although the accuracy of time data is often diminished by unanticipated interruptions, personal breaks and general productivity variability efforts should be made to include these factors in estimates of the total time to complete the PI. Conversely, very long time frames that are obviously anomalous should not be included in the assessment. It was assumed that any point taking more than 20 min to complete (as indicated by the time stamp) was not a representative observation and was disregarded in the time analysis (1.2% of observations). The average time per PI point was calculated and multiplied by various salary rates to obtain the cost per point for several workforce cost scenarios. Due to the same set of field plots being used throughout the analysis, the field data collection cost is the same under both sampling paradigms. All other costs associated with forest inventory implementation and administration were considered to be the same under both sampling methods.

Results/discussion

As expected, the county estimates for forestland area and total biomass were slightly dissimilar between the PS and DSPS designs (Table 1)—primarily attributable to unequal stratum weights. However, the precision of the estimates as indicated by the sampling errors differed substantially, with biomass being less precise than forestland area. In all cases, the sampling error increases as the PI point intensity decreases from the original intensity of 25 points per plot (PI 100%). A notable outcome was that the PI effort could be reduced to 25% (PI 25%—6 points per plot) and the DSPS estimates would still be more precise that PS for estimates of forest area. The PI effort could similarly be reduced by 50% (PI 50%) while still outperforming PS for estimates of total biomass. Using the DSPS design and obtaining PI information only where the sample plots occur resulted in considerable decreases in precision compared to PS.

It is also interesting to note the effect of reducing the PI effort has only marginal effects on the sampling errors. For example, the SE% only increases by about 10% when going from PI 100% to PI 25%. The degradation of the sampling error increases dramatically between PI 25% (6 points per plot) and PI plots (1 point at the plot location). Further investigation into relatively small PI efforts showed increasing rates of degradation as efforts dropped below PI 25% (Fig. 3). Generally, PI efforts below PI 25% are not recommended due to the precision loss relative to more intensive efforts and the fact that the PS design becomes more favorable when the PI effort is minimal. It is also shown that if the number of PI points is equal to the number of sample plots, there is considerable disadvantage to co-locating the PI points and the plot locations (PI plot). Presumably, little new information is obtained in the co-location scenario as compared to having the PI points positioned elsewhere.

A factor affecting the outcomes is the temporal frame of the remote sensing information used in the stratification. In this study, the NLCD map layer was from 2011, the NAIP imagery from 2015, and the plot data collected over the 2011–2016 period. If one considers the midpoint of the data collection effort to be 2014, the NAIP imagery is on average closer to the plot observation time than the NLCD map. Thus, the NAIP stratification may have some advantage in this regard, depending on how much the populations have changed between 2011 and 2015. This situation points to the temporal benefits of using NAIP in comparison with NLCD for two reasons. First, NAIP is updated on a 2–3-year interval, whereas NLCD data are presented at 5-year increments. Second, availability of NAIP imagery occurs soon after the photography is completed; but the NLCD product usually takes several years beyond the base date to be completed.

Table 1 Estimates and sampling errors for total forestland area (ha) and tree biomass (tonnes) for 3 counties using the PS and DSPS designs

County	Design	Stratification	Points	Biomass (tonnes)	SE%	Forest area (ha)	SE%
Schuylkill	DSPS	PI 100%	2099	17,339,268	7.05	124,311	3.65
	DSPS	PI 75%	1575	17,351,687	7.11	124,382	3.73
	DSPS	PI 50%	1050	17,381,081	7.21	124,555	3.88
	DSPS	PI 25%	525	17,415,935	7.51	124,744	4.28
	DSPS	PI Plots	71	18,119,576	10.39	128,754	7.62
	PS	NLCD Map		18,157,560	7.25	130,383	4.58
Lehigh	DSPS	PI 100%	925	5,412,296	19.12	24,276	14.01
	DSPS	PI 75%	694	5,409,749	19.25	24,264	14.20
	DSPS	PI 50%	463	5,426,299	19.46	24,338	14.51
	DSPS	PI 25%	231	5,430,023	20.09	24,356	15.38
	DSPS	PI Plots	34	5,492,805	26.75	24,640	23.86
	PS	NLCD Map		5,479,217	23.19	21,995	16.98
Berks	DSPS	PI 100%	2350	14,374,125	9.02	78,178	6.23
	DSPS	PI 75%	1763	14,355,313	9.13	78,094	6.38
	DSPS	PI 50%	1175	14,324,136	9.33	77,955	6.66
	DSPS	PI 25%	588	14,228,979	9.87	77,508	7.38
	DSPS	PI Plots	85	13,069,231	15.34	72,205	13.86
	PS	NLCD Map		12,620,458	9.68	69,638	7.83

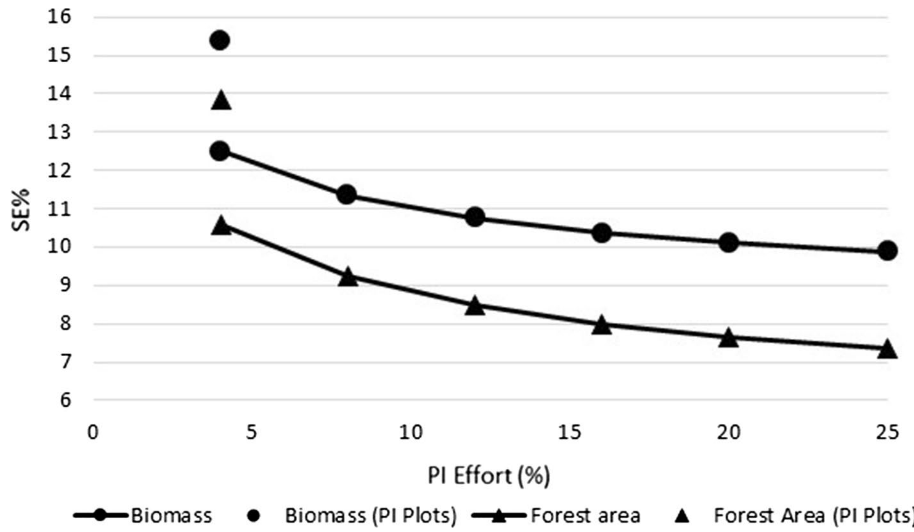

Fig. 3 Sampling error (SE%) for low levels of PI effort using the DSPS design in Berks County, PA. There are two SE% given for each attribute at the PI effort of 4%—one reflecting the PI being done at the plot locations (PI Plots) and the other indicating the same level of PI effort at nonplot PI points

Table 2 Cost per PI plot (US $) comparison matrix for several levels of salary cost and interpretation productivity rates

Salary (US $/h)	Time per points (s)		
	11	23	34
50	$0.153	$0.319	$0.472
35	$0.107	$0.224	$0.331
20	$0.061	$0.128	$0.189
10	$0.031	$0.064	$0.094

Table 3 Additional cost per inventory plot for DSPS at various levels of PI effort and interpretation time per point (assumes salary of US $35/h)

Time per points (s)			
PI effort (%)	11	23	34
PI 100	$2.67	$5.59	$8.26
PI 75	$2.01	$4.19	$6.20
PI 50	$1.34	$2.80	$4.13
PI 25	$0.67	$1.40	$2.07

More generally, the closer the remote sensing data are to the plot measurement date, it is more likely the stratification effectiveness will increase (McRoberts et al. 2016).

Estimating the cost of the PI stratification was based on average time per PI point. The mean value was 23 s and the corresponding cost to perform the work was approximately US $35/h. Thus, the cost per PI point was approximately US $0.224 and the entire cost for the PI work across all 3 counties was nearly US $1200. It is recognized that a range of other cost scenarios may be encountered—such as differing rates of productivity and/or staff costs. Table 2 provides comparative information that may help inform costs of PI work in other situations. Realized production rates are subject to numerous factors such as landscape heterogeneity, imagery resolution and/or photo-interpreter experience. The data collected in this study showed no discernable correlation between the photo-interpretation time and canopy cover percent at the PI point.

Of considerable importance is the evaluation of the cost of the PI work in the context of overall inventory costs and precision gains. Because the PI effort is expressed on a PI points per ground plot basis, the additional expense can be given as an increase in cost per plot. Table 3 shows the additional cost per plot is very small, regardless of PI effort and interpreter efficiency. Given an average cost of US $600 to complete a ground plot, a 100% PI effort and interpreter rate of 34 s/point would result in an increase in cost per plot of about 1.4%, although our data suggest the cost is more likely to be slightly less than 1.0% (23 s/point). If the PI effort was reduced to 50%, DSPS would still provide superior precision than PS with an additional cost incursion of less than 0.5% per plot.

An alternative method of assessing cost differences between the DSPS and PS designs is to determine the additional number of field plots required to attain a specified level of precision. For the purposes of comparison, the additional number of plots needed in the PS design to attain the precision of DSPS was examined for Berks County. To obtain DSPS precision at the PI 100% level, the PS design would require 15% and 58% more plots for estimates of total biomass and forestland area, respectively (Table 4). Reducing the effort to PI 50% still shows DSPS to be more cost-effective (0.5% increase) than PS (8% and 38% increase, respectively). At PI 25%, the PS design becomes more cost-effective than DSPS for total biomass, but is still considerably more costly for forestland area. Essentially, as long as the percent sample size increase for PS exceeds the percent

Table 4 Percent change in sample size needed under the PS design to obtain the same precision obtained from the DSPS design for Berks County

PI effort (%)	Sample size change	
	Biomass (%)	Forest area (%)
PI 100	15	58
PI 75	12	51
PI 50	8	38
PI 25	−4	13

cost increase for PI work, the DSPS design is less expensive to implement for the same level of precision.

The cost evaluations were based on a number of factors specific to the FIA program in the study area. Most notably, the sample plot size and configuration, number of attributes to be measured, travel time (depends on road density, topography, etc.), and number of crew members and their salary rates. As such, cost assessments for other inventories should be conducted with respect to the specific circumstances encountered. Similarly, the cost of performing the PI work should be evaluated, e.g., as shown in Tables 2 and 3. This reassessment of the PI is particularly important if there is a different configuration and/or number of grid points within each PI point than was used here. Depending on the outcomes of these assessments, DSPS may become more or less favorable than PS as shown in this analysis.

No cost was ascribed to development of the stratification for PS based on the NCLD canopy cover map. It was assumed that software would be generally available for a number of tasks related to remote sensing analysis, and thus the software cost should not be assigned to this specific task, although there may be situations where the software is primarily obtained and used for stratification purposes, in which the cost should then be accounted for. Some amount of staff effort is needed to perform the calculations from the NLCD map; while comparatively, staff effort would also be incurred to organize the imagery for the DSPS stratification. In this study, it was assumed the staff time was similar in either case and generally negligible in the context of the other costs. This may not be the case in all situations and any notable differences between the two approaches should be accounted for in the cost assessment.

Lastly, for existing forest inventories where many features such as sample sizes, plot locations and wall-to-wall PS methods are already in place, changing to DSPS represents an increase in the cost of the inventory. Even though only a small relative cost increase is incurred and the advantages in precision of estimates can be substantial, it may be that additional costs simply cannot be incurred, i.e., the inventory was designed and structured within the funding limitations. In such cases, DSPS may not be a viable alternative, although perhaps efficiencies in other parts of the inventory operation can be found and savings redirected at PI work. For new inventory endeavors, DSPS should be considered as a viable alternative when evaluating cost:precision scenarios during the inventory design planning phase.

Conclusion

Application of the double sampling for stratification paradigm to an existing forest inventory required conceptualization of the DSPS design, and in particular, the development of an appropriate variance estimator that accounts for both random stratum sample sizes and estimated stratum weights. This allowed for valid comparisons in precision between the PS and DSPS designs. The initial DSPS effort of 25 PI points per sample plot produced better precision than PS for estimates of total biomass and forestland area; however, even less intensive efforts (PI 50% for biomass; PI 25% for forestland area) could be conducted while still maintaining more precise estimates from DSPS. The results also suggested that 25 PI points per plot may be excessive, in that the precision decreases only by about 10% when 6 PI points per plot are used; however, these outcomes may differ for other populations and ultimately the PI effort needs to be carefully examined for efficiency:precision relationships that exist for a given population of interest. It is also worth noting the DSPS design performs remarkably well despite the vastly smaller number of information elements available for the stratification—suggesting the canopy cover accuracy of the PI work is substantially higher than that obtained from classified digital maps.

A primary disadvantage of conducting PI work is the additional cost. However, depending on the PI effort, the methods used in this study suggest the cost increase per sample plot is on the order of 0.5–1.0%. In comparison with obtaining the same precision under the PS design, the DSPS design is considerably more cost-effective as long as the PI effort is maintained at or above PI 50% for biomass and PI 25% for forestland area. DSPS becomes more costly when the PI effort is too small and the resulting precision is near or less than that from the PS design. The cost-effectiveness of any stratified forest inventory design needs to be evaluated considering all factors pertaining to conducting both the sample plot measurements and the cost of developing the stratification scheme for estimation. Thus, the results of this study should be not construed as being applicable to all situations, but the methods and outcomes may provide a framework and guidance for conducting cost:precision assessments to determine the viability of the DSPS design in other environments.

Acknowledgements The authors are grateful to the associate editor, Dr. Tim Gregoire, and an anonymous reviewer for their insightful comments that resulted in considerable improvement to the paper.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bechtold WA, Scott CT (2005) The forest inventory and analysis plot design. In: Bechtold WA, Patterson PL (eds) The enhanced forest inventory and analysis program—national sampling design and estimation procedures. U.S. Department of Agriculture, Forest Service, Southern Research Station, Gen. Tech. Rep. SRS-80, Asheville, pp 37–52

Bickford CA (1952) The sampling design used in the forest survey of the Northeast. J For 50(4):290–293

Brown JP, Westfall JA (2012) An evaluation of the properties of the variance estimator used by FIA. In: McWilliams W, Roesch FA (eds) Monitoring across borders: 2010 joint meeting of the forest inventory and analysis (FIA) symposium and the southern mensurationists. Monitoring Across Borders: 2010 Joint Meeting of the Forest Inventory and Analysis (FIA) Symposium and the Southern Mensurationists, Asheville, pp 53–58

Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York

De Gruijter J, Brus DJ, Bierkens MF, Knotters M (2006) Sampling for natural resource monitoring. Springer, Berlin

De Vries PG (1986) Sampling theory for forest inventory: a teach-yourself course. Springer, Berlin

Durbin J (1969) Inferential aspects of the randomness of sample size in survey sampling. In: Johnson NL, Smith H Jr (eds) New developments in survey sampling. Wiley, New York, pp 629–651

Frayer WE, Furnival GM (1999) Forest survey sampling designs: a history. J For 97(12):4–10

Gregoire TG, Ringvall AH, Ståhl G, Næsset E (2016) Conditioning post-stratified inference following two-stage, equal-probability sampling. Environ Ecol Stat 23(1):141–154

Hansen MH, Wendt DG (2000) Using classified landsat thematic mapper data for stratification in a statewide forest inventory. In: McRoberts RE, Reams GA, Van Deusen PC (eds) Proceedings of the first annual forest inventory and analysis symposium. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station, General Technical Report. NC-213, pp 20–27

Holt D, Smith T (1979) Post stratification. J R Stat Soc Series A (General) 142(1):33–46

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold ND, Wickham JD, Megown K (2015) Completion of the 2011 national land cover database for the conterminous United States-representing a decade of land cover change information. Photogram Eng Rem Sens 81(5):345–354

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

Köhl M, Magnussen SS, Marchetti M (2006) Sampling methods, remote sensing and GIS multiresource forest inventory. Springer, Berlin

Köhl M, Lister A, Scott CT, Baldauf T, Plugge D (2011) Implications of sampling design and sample size for national carbon accounting systems. Carbon Balance Manag 6(1):10

Lam TY, Kleinn C, Coenradie B (2011) Double sampling for stratification for the monitoring of sparse tree populations: the example of populus euphratica Oliv. forests at the lower reaches of Tarim River, Southern Xinjiang, China. Environ Monit Assess 175(1):45–61

Lister A, Scott C (2009) Use of space-filling curves to select sample locations in natural resource monitoring studies. Environ Monit Assess 149(1–4):71–80

McRoberts RE, Nelson MD, Wendt DG (2002) Stratified estimation of forest area using satellite imagery, inventory data, and the k-nearest neighbors technique. Remote Sens Environ 82:457–468

McRoberts RE, Næsset E, Gobakken T (2013) Inference for lidar-assisted estimation of forest growing stock volume. Remote Sens Environ 128:268–275

McRoberts RE, Næsset E, Gobakken T (2016) The effects of temporal differences between map and ground data on map-assisted estimates of forest area and biomass. Ann For Sci 73:839–847

Reams GA, Smith WD, Hansen MH, Bechtold WA, Roesch FA, Moisen GG (2005) The forest inventory and analysis sampling frame. In: Bechtold WA, Patterson PL (eds) The enhanced forest inventory and analysis program—national sampling design and estimation procedures. U.S. Department of Agriculture, Forest Service, Southern Research Station, General Technical Report. SRS-80, Asheville, pp 21–36

Saborowski J, Cancino J (2007) About the benefits of poststratification in forest inventories. J For Sci 53(4):139–148

Scott CT, Bechtold WA, Reams GA, Smith WD, Westfall JA, Hansen MH, Moisen GG (2005) Sample based estimators used by forest inventory and analysis national information management system. In: Bechtold WA, Patterson PL (eds) The enhanced forest inventory and analysis program—national sampling design and estimation procedures. U.S. Department of Agriculture, Forest Service, Southern Research Station, General Technical Report. SRS-80, Asheville, pp 43–67

Stephan FF (1945) The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate. Ann Math Stat 16(1):50–61

Tomppo E, Olsson H, Ståhl G, Nilsson M, Hagner O, Katila M (2008) Combining national forest inventory field plots and remote sensing data for forest databases. Remote Sens Environ 112(5):1982–1999

U.S. Forest Service (2012) National core field guide: version 6.0. Vol. 1: field data collection procedures for phase 2 plots. https://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2013/Core%20FIA%20P2%20field%20guide_6-0_6_27_2013.pdf. Accessed 15 June 2018

USDA-FSA-APFO (2016) National agriculture imagery program (NAIP) 2015 for Pennsylvania. http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=3183. Accessed 7 March 2018

Valliant R (1993) Poststratification and conditional variance estimation. J Am Stat Assoc 88(421):89–96

Wayman JP, Wynne RH, Scrivani JA, Reams GA (2001) Landsat TM-based forest area estimation using iterative guided spectral class rejection. Photogram Eng Remote Sens 67(10):1155–1166

Westfall JA, Patterson PL, Coulston JW (2011) Post-stratified estimation: with-in strata and total sample size recommendations. Can J For Res 41:1130–1139

Westfall JA, Lister AJ, Scott CT (2016) Precision and cost considerations for two-stage sampling in a panelized forest inventory design. Environ Monit Assess 188:11