Double sampling for stratification in periodic inventories—Infinite population approach

Joachim Saborowski [a,b,*], Almuth Marx [c], Jürgen Nagel [d], Thomas Böckmann [e]

[a] Department of Ecoinformatics, Biometrics and Forest Growth, University of Göttingen, Büsgenweg 4, D-37077 Göttingen, Germany
[b] Department of Ecosystem Modelling, University of Göttingen, Büsgenweg 4, D-37077 Göttingen, Germany
[c] Freilandstr. 12 a, D-90482 Nürnberg, Germany
[d] Northwest German Forest Research Station, Grätzelstr. 2, D-37079 Göttingen, Germany
[e] Lower Saxony Forest Planning Office, Forstweg 1a, D-38302 Wolfenbüttel, Germany

ARTICLE INFO

ABSTRACT

Double sampling for stratification is a sampling design that is widely used for forest and other resource inventories in forest ecosystems. It is shown that this sampling design can be adapted to repeated inventories including estimators of net change, even for non-proportional allocation of second-phase units and periodically updated stratification. The method accounts for the transition of sampling units among strata. Moreover, it may outperform classical single phase designs if sample plots are appropriately allocated to strata with respect to predefined target variables, here: volume per ha of bigger trees of the main tree species. The latter requires a clear definition of predominant aims of the inventory and an appropriate optimization method. Access to inventory data of a state forest district from two occasions allowed for an optimization of the design based on the first occasion, which proved to be still advantageous on the following occasion. Estimators are developed under the infinite population approach, which is generally deemed more appropriate for forest inventories.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Stratified sampling and in particular double or 2-phase sampling for stratification (2SS) are basic techniques treated in all standard textbooks on theory and application of sampling techniques (e.g. Cochran, 1977; Thompson, 1992; Särndal et al., 2003), especially in those focussing on forest inventory designs (de Vries, 1986; Schreuder et al., 1993; Gregoire and Valentine, 2008; Mandallaz, 2008). It is mostly used together with remote sensing techniques (aerial photographs, satellite images), which usually allow for cheap selection and stratification of first-phase units. 2SS can easily be extended to multiphase designs for stratification. A comprehensive collection of formulas and examples for two, three and four phases can be found in Frayer (1979). Often, double sampling is used in kind of a degenerate version where a population is split into only two strata, one that consists of units which have only zero values and another one with values assumed non-zero (Singh and Singh, 1965b), e.g. non-forest and forest units.

There are numerous applications of 2SS in forest inventory, predominantly in countries with large forest areas such as Canada, the United States or the former Soviet Union (Gabler and Schadauer, 2007). In the United States, the Forest Inventory and Analysis (FIA) programme is conducted since the 1930s, where about 2,500,000 photo-interpreted points and 120,000 fixed area plots on forested land were measured to monitor forest resources (Williams, 2001). On a smaller but still large scale are the county timber inventories in the State of Washington, which were used by Mac Lean (1972) to compare the efficiency of different stratifications and allocations of sampling units.

Whereas 2SS was already well known much earlier, Singh and Singh (1965a) and Rao (1973) published subsampling procedures (second phase) that are mathematically sound and free of the inconsistency which arose from the assumption that the second-phase sample size n_h within stratum h is fixed, an assumption that was usually made in earlier literature but contradicted the sampling procedure as carried out in practice. 2SS can also be used on successive occasions, which is an interesting application from the viewpoint of forest inventory, because it allows for Sampling with Partial Replacement (SPR) within the strata. Pure SPR was introduced for Continuous Forest Inventory (CFI) by Ware and Cunia (1962). Bickford et al. (1963) proposed a combination of a two occasions SPR and 2SS with a new and independent first-phase selection on the second occasion, but did not provide an estimator of change.

* Corresponding author at: Department of Ecoinformatics, Biometrics and Forest Growth, University of Göttingen, Büsgenweg 4, D-37077 Göttingen, Germany. Tel.: +49 551 393450; fax: +49 551 393465.
E-mail address: jsaboro@gwdg.de (J. Saborowski).

Table 1
$L = 8$ Strata of first-phase sample plots (DEC: deciduous trees dominating, CON: coniferous trees dominating).

		Age classes			
		0–40	>40–80	>80–120	>120
Dominating species group	Deciduous	DEC1 ($h=1$)	DEC2 ($h=2$)	DEC3 ($h=3$)	DEC4 ($h=4$)
	Coniferous	CON1 ($h=5$)	CON2 ($h=6$)	CON3 ($h=7$)	CON4 ($h=8$)

It was not until 1994 that Scott and Köhl (1994) extended that design to three occasions and the estimation of net changes. Singh and Singh (1965b) also used 2SS on successive occasions for the estimation of coconut production in Assam. They assumed a static stratification, i.e. no shifts of units from one stratum to another in the course of time, as well as no addition of units to or removal of units from a stratum, a rather unrealistic assumption in the case of CFIs.

Such changes of the structure of a population with respect to the strata are a crucial point if 2SS is applied on two or several occasions. That problem was also mentioned by Scott and Köhl (1994), who noted that the effectiveness of stratification can be degraded when the new second-phase sample on the second occasion is selected from the first-phase sample of the first occasion. They suggested an alternative method based on an independent new first-phase sample on occasion 2 with succeeding stratification of both the new first-phase sample and, additionally, all second-phase units measured at occasion 1. Additionally, they assume proportional allocation of the second-phase sample, what is optimal if variances within strata are equal and costs are not considered.

Based on the preceding considerations, we propose a 2SS design for repeated inventories which allows for non-proportional allocation of second-phase units and thereby for optimization of second-phase sampling proportions. The stratification of first-phase units chosen on the first occasion can be updated on every new occasion. We will show how one can solve the problem of shifting units mentioned above, and we will derive unbiased estimators of means and their variances under the infinite population approach. We do not consider SPR within strata. The method is demonstrated using data from a forest district inventory, but it is not restricted to forest inventories.

2. Data base

In 1999 a 2SS design was applied in the Saupark State Forest district of Lower Saxony (Germany) using first-phase sample points on a 100 m × 100 m grid intersected with the forest area. At each grid point a virtual ground plot of 13 m radius was assigned to one of $L = 8$ strata (Table 1) by interpretation of CIR aerial images. The strata are defined by age class (4 intervals of 40 years) and dominating tree species group (two classes: deciduous and coniferous). The idea behind that stratification was that age and species groups are closely related to volume, and 4 age classes as well as 2 dominating tree species groups can easily be distinguished by interpretation of CIR images at low costs. The number of age strata should be low, because shifts of sampling units from one stratum to another were expected to be a source of trouble in repeated inventories and should be kept on a low level. Within the usual inventory period of 10 years, roughly 25% of the sampling units of a stratum can be expected to move naturally to the next older 40-year age class or to the youngest. Last but not least, strata should be large, because an efficient allocation of second-phase ground plots derived from current inventory data should at least approximately hold for the repeated inventory. That requires that variances within strata do not change remarkably within one inventory period. The aforementioned arguments are to be considered particularly in more extensive inventories using a much larger number of strata.

Here and in the following, "volume" means volume of standing, alive trees regarding wood of at least 7 cm diameter at breast height (dbh).

In the second phase, a predefined proportion of first-phase units (grid points) was selected, where 2 concentric circles of 13 m (dbh ≥ 30 cm) and 6 m (7 cm ≤ dbh < 30 cm) were established as terrestrial sample plots. Second-phase sample plots were systematically selected from a one-dimensional list of all first-phase units in a stratum. We note that the 2SS method described in this article will instead assume simple random sampling in the second phase and uniform point sampling in the first phase.

The proportions of second-phase sample sizes within strata related to the respective first-phase sample sizes tend to be larger in higher age classes (Table 2a) and are larger for strata dominated by coniferous trees. As will be shown later, this was not optimal for the estimation of volume per hectare of higher diameter classes of the main tree species. The total number of terrestrial plots or second-phase units was 1644. In Table 2 we report proportions normalized to a second-phase sample size of 1000, because it allows to adapt those proportions easily to other second-phase sample sizes. If e.g. higher precision of estimates is required, the second-phase sample size can be increased leading to proportionally larger sampling proportions.

End of 2008, 9 years after the first inventory, all second-phase sampling units of the first occasion were measured a second time. We note that the first-phase stratification was not updated before remeasurement on that second occasion. Only for the purpose of this study, the first-phase stratification was updated later in January 2010 based on CIR images from September 2008. In the following chapters, we will refer to 2008 as the year of the second occasion.

Table 3 shows a cross-tabulation of 6803 first-phase units stratified on both occasions. As could be expected, most of these units (74%) remained in the stratum which they were assigned to in 1999 (main diagonal of Table 3), or they moved to the next older stratum, or, in case they belonged to the oldest stratum in 1999, to the youngest. Nearly all other transitions were also observed, yet to a lower degree. They can be explained by removal of trees from plots changing the dominating tree species in mixed stands or changing the age class in two- or multilayered stands, but evidently also by erroneous interpretation either in 1999 or in 2008. Such errors may particularly occur if plot centres are not exactly located in the aerial image.

Table 2
Second-phase sampling proportions $v_{h,\text{norm}}$ per stratum applied in 1999, and optimum proportions based on variance estimates from the 1999 data, normalized with respect to $n = 1000$. For a total second-phase sample size of $n = 1644$ choose $v_h = n_h/n'_h = v_{h,\text{norm}} \cdot 1.644$, where n'_h is the number of first-phase units in stratum h.

		Age classes			
		0–40	>40–80	>80–120	>120
(a) Applied 1999	Deciduous	0.11	0.07	0.09	0.13
	Coniferous	0.19	0.28	0.24	0.26
(b) Optimized	Deciduous	0.06	0.13	0.13	0.21
	Coniferous	0.08	0.10	0.30	0.40

Table 3
Cross-tabulation of $n' = 6803$ first-phase sampling units stratified using CIR images from 1999 and 2008. Shaded fields represent the transitions expected by aging and by harvesting and subsequent reforestation with previous dominating tree species group.

		Stratum l in 1999								Total
		DEC1	DEC2	DEC3	DEC4	CON1	CON2	CON3	CON4	
Stratum h in 2008	DEC1	563	18	7	54	19	17	25	6	709
	DEC2	260	857	12	26	35	66	7	0	1263
	DEC3	7	220	746	36	5	30	26	1	1071
	DEC4	67	29	191	1535	5	10	11	5	1853
	CON1	18	5	0	2	125	2	7	2	161
	CON2	13	35	9	11	283	1114	6	5	1476
	CON3	8	3	1	3	7	94	109	0	225
	CON4	1	1	0	7	2	2	23	9	45
	Total	937	1168	966	1674	481	1335	214	28	6803

3. Infinite population approach to the 2SS sampling design

We assume a classical double sampling design for stratification with random sampling in both phases similar to that defined in Cochran (1977) or Särndal et al. (2003), except that we consider an infinite population approach in the first phase. In the second phase of a 2SS design, sampling is done from a finite number of first-phase units and can be described by the classical finite population approach as usual. We note that Mandallaz (2008) provided an infinite population approach for a design called one-stage double sampling for stratification which does not fit to the design applied here, because second-phase units are randomly chosen from the entire population of first-phase sampling points in his approach, and not independently within strata and with different predefined sampling proportions. In this sense his design represents a post-stratification as is also pointed out there.

Our notation is a compromise of those used in Cochran (1977) and Mandallaz (2008). Let a population of N trees lie in a forest area F of surface area $\lambda(F)$, stratified into L disjoined strata $F_h(h = 1, \ldots, L)$ with $\lambda(F) = \sum \lambda(F_h)$ and $N = \sum N_h$. It is assumed that the first-phase sample σ', a point sample of size n', is independently and uniformly distributed within F, and each sample point $x \in \sigma'$ is assigned to one of the L strata using auxiliary variables. Let n'_h be the number of first-phase sample points $x \in \sigma_h$ with $n' = \sum n'_h$. In this approach, elements of the target population (trees) and sampling units (points) differ, while Cochran (1977) uses N and N_h for the finite number of population elements, which are at the same time the sampling units.

In the second phase, a random subsample σ_h of size $n_h = \nu_h n'_h$ be drawn without replacement from the n'_h first-phase sample points according to a predefined proportion ν_h. The second-phase sample sizes $\nu_h n'_h$ may be non-integer and must be rounded to the next integer value. At each second-phase point x the local density (Mandallaz, 2008)

$$Y(x) = \frac{1}{\lambda(F_h)} \sum_{i=1}^{N_h} \frac{I_i(x) Y_i}{\pi_i}, \tag{1}$$

with Y_i the response variable of tree i, is provided. $\pi_i = P(I_i(x) = 1)$ is the inclusion probability of tree i, and the tree is included in the sample if it lies in the sample plot at sample point x. For concentric circles, which are often used in forest inventories, it holds $\pi_i = \lambda(K_i \cap F_h)/\lambda(F_h)$, where K_i is a circle centred at tree i with a radius depending on the tree diameter, and

$$I_i(x) = \begin{cases} 1 : x \in K_i \\ 0 : x \notin K_i \end{cases}.$$

The target parameters of the population are

$$\bar{Y} = \frac{1}{\lambda(F)} \sum_{i=1}^{N} Y_i \quad \bar{Y}_h = \frac{1}{\lambda(F_h)} \sum_{i=1}^{N_h} Y_i.$$

Since the local density is constructed such that its expectation

$$E\,Y(x) = \frac{1}{\lambda(F_h)} \int_{F_h} Y(x) dx$$

for a uniformly distributed random point $x \in F_h$ equals the population parameter $\bar{Y}_h$, unbiased estimators of $\bar{Y}$ and $\bar{Y}_h$ are

$$\hat{\bar{Y}} = \sum_{h=1}^{L} w_h \hat{\bar{Y}}_h \quad \hat{\bar{Y}}_h = \frac{1}{n_h} \sum_{x \in \sigma_h} Y(x), \tag{2}$$

respectively, with weights $w_h = n'_h/n'$ and $E w_h = \lambda(F_h)/\lambda(F) =: W_h$. Deviating from the notation of Mandallaz (2008), we use $\hat{\bar{Y}}$ and $\hat{\bar{Y}}_h$ instead of $\hat{Y}$ and $\hat{Y}_h$.

The variance of $\hat{\bar{Y}}$ is given by (Appendix A)

$$V(\hat{\bar{Y}}) = \frac{1}{n'} \left(\sum_{h=1}^{L} \frac{W_h S_h^2}{\nu_h} + \sum_{h=1}^{L} W_h (\bar{Y}_h - \bar{Y})^2 \right), \tag{3}$$

where

$$S_h^2 = \frac{1}{\lambda(F_h)} \int_{F_h} (Y(x) - \bar{Y}_h)^2 dx$$

This formula is similar to Cochran's approximate formula (12.14) and structurally equal for large population size N and $g' \approx 1$ (Cochran, 1977), but it is closer to forest inventory practice because of the more appropriate definitions of S_h^2 and W_h. Moreover, it will lead to a much simpler unbiased variance estimator than that given by Cochran's (12.25) for the finite population approach.

Cochran (1977) points out that a sample copy of (12.14) would be an almost unbiased estimator in almost all applications. Here, the sample copy of (3) is

$$\hat{V}(\hat{\bar{Y}}) = \frac{1}{n'} \left(\sum_{h=1}^{L} \frac{w_h s_h^2}{\nu_h} + \sum_{h=1}^{L} w_h (\hat{\bar{Y}}_h - \hat{\bar{Y}})^2 \right).$$

$$s_h^2 = \frac{1}{n_h - 1} \sum_{x \in \sigma_h} (Y(x) - \hat{\bar{Y}}_h)^2, \tag{4}$$

but a strictly unbiased estimator of (3) is given by

$$
\begin{aligned}
\hat{V}(\hat{\bar{Y}}) &= \frac{1}{n'-1}\left(\sum_{h=1}^{L}\frac{w_h s_h^2}{v_h} + \sum_{h=1}^{L} w_h\left(\hat{\bar{Y}}_h - \hat{\bar{Y}}\right)^2 - \sum_{h=1}^{L}\frac{s_h^2}{n' v_h}\right) \\
&= \frac{1}{n'-1}\left(\sum_{h=1}^{L}\frac{n_h'-1}{n'}\frac{s_h^2}{v_h} + \sum_{h=1}^{L} w_h\left(\hat{\bar{Y}}_h - \hat{\bar{Y}}\right)^2\right)
\end{aligned}
\qquad (5)
$$

which predominantly differs from (4) by the factor $(n_h'-1)/n'$ instead of $w_h = n_h'/n'$, or equivalently by the third sum within the parentheses. That term will usually be small for large n'. Again, we can state a structural similarity with the finite population approach of Cochran (1977) comparing his unbiased variance estimator (12.25) with (5) under the assumptions $g' \approx 1$ and N large. See Appendix A for a proof of (5).

4. Sampling procedure on the second occasion

At the second occasion the first-phase stratification of all sampling units is assumed to be repeated leading to an updated allocation of first-phase units to the L strata:

$$
m' = \sum_{h=1}^{L} m_h'
$$

Compared to the first occasion, a certain number of units will have changed their stratum, because they have passed an age threshold, or the dominant tree species group has changed from deciduous to coniferous or vice versa, or by removal of a stand layer. Consequently, it cannot be assumed that $m_h' = n_h'$ for the L strata, and the proportion of second-phase units among the first-phase units in a stratum will also change, even if $m' = n'$.

Consequently, the m_h' first-phase units assigned to stratum h on the second occasion may comprise units which had been assigned to any of the strata on the first occasion, i.e.

$$
m_h' = \sum_{l=1}^{L} m_{hl}'
$$

where m_{hl}' represents the number of first-phase units which moved from stratum l to stratum h, and in particular m_{hh}' the number of units which did not change strata. Since the expected number of second-phase plots among each group of m_{hl}' first-phase plots is $v_l m_{hl}'$, with first occasion sampling proportions v_l usually varying among those groups, one has to select additional second-phase plots or to delete existing plots in order to achieve any required sampling proportion $\mu_h = m_h/m_h'$, even if $\mu_h = v_h$. That

additional selection or deletion has to be done groupwise to avoid over- or underrepresentation of first-phase units from different original strata and yields second-phase sample sizes $m_{hl} = \mu_h m_{hl}'$ with $m_h = \sum_{l=1}^{L} m_{hl}$. Given the first-phase stratification the m_{hl}' are fixed and known and form substrata of the m_h' units of stratum h. The m_{hl} second-phase plots may comprise only existing plots that moved from l to h, occasionally randomly reduced to meet $m_{hl} = \mu_h m_{hl}'$, or they may comprise existing plots as well as additional randomly selected plots newly established on the second occasion if the number of existing plots was smaller than $\mu_h m_{hl}'$. According to this procedure all m_{hl}' first-phase units have equal chances of being selected as second-phase plots leading to simple random sampling without replacement from the m_{hl}' first-phase units, or in other words to equal inclusion probabilities $\mu_h = m_{hl}/m_{hl}'$ ($l = 1, \ldots, L$) within all substrata, i.e. proportional allocation.

The estimator (2) for 2SS (the target variable Y is replaced by Z indicating the volume per ha on the second occasion)

$$
\hat{\bar{Z}} = \sum_{h=1}^{L}\frac{m_h'}{m'}\hat{\bar{Z}}_h
$$

is adapted to that stratified subsampling procedure of the second phase using

$$
\hat{\bar{Z}}_{h,\mathrm{prop}} = \sum_{l=1}^{L}\frac{m_{hl}'}{m_h'}\hat{\bar{Z}}_{hl} = \frac{1}{m_h}\sum_{l=1}^{L}\sum_{j=1}^{m_{hl}} z_{hlj}
$$

instead of $\hat{\bar{Z}}_h$. $\hat{\bar{Z}}_{h,\mathrm{prop}}$ is the well known estimator (e.g. Cochran, 1977) for stratified random sampling with proportional allocation. For notational convenience, $z_{hlj} = Z(x_{hlj})$ be the local density at a sample point randomly selected from all first-phase sample points which have moved from stratum l to stratum h. Thus,

$$
\hat{\bar{Z}}_{\mathrm{prop}} = \sum_{h=1}^{L}\frac{m_h'}{m'}\hat{\bar{Z}}_{h,\mathrm{prop}} = \sum_{h=1}^{L}\left(\frac{m_h'}{m'}\cdot\frac{1}{m_h}\sum_{l=1}^{L}\sum_{j=1}^{m_{hl}} z_{hlj}\right)
\qquad (6)
$$

is an unbiased estimator of $\bar{Z}$ with the true variance

$$
V(\hat{\bar{Z}}_{prop}) = \frac{1}{m'}\sum_{h=1}^{L} W_h^{(2)} S_{zh}^2 + \frac{1}{m'}\sum_{h=1}^{L} W_h^{(2)}\left(\bar{Z}_h - \bar{Z}\right)^2 + E_1\left(\frac{1}{m'}\sum_{h=1}^{L}\frac{m_h'}{m'}\left(\frac{1}{\mu_h}-1\right)\sum_{l=1}^{L}\frac{m_{hl}'}{m_h'} s_{zhl}'^2\right)
$$

$$
S_{zh}^2 = \sum_{l=1}^{L} W_{hl}^{(2)} S_{zhl}^2 + \sum_{l=1}^{L} W_{hl}^{(2)}\left(\bar{Z}_{hl} - \bar{Z}_h\right)^2
\qquad (7)
$$

and the sample copy as a variance estimator

$$
\begin{aligned}
\hat{V}(\hat{\bar{Z}}_{\mathrm{prop}}) &= \frac{1}{m'}\sum_{h=1}^{L}\frac{m_h'}{m'}\hat{S}_{zh}^2 + \frac{1}{m'}\sum_{h=1}^{L}\frac{m_h'}{m'}\left(\hat{\bar{Z}}_{h,\mathrm{prop}} - \hat{\bar{Z}}_{\mathrm{prop}}\right)^2 + \frac{1}{m'}\sum_{h=1}^{L}\frac{m_h'}{m'}\left(\frac{1}{\mu_h}-1\right)\sum_{l=1}^{L}\frac{m_{hl}'}{m_h'} s_{zhl}^2 \\
&= \frac{1}{m'}\left(\sum_{h=1}^{L}\frac{w_h^{(2)}}{\mu_h}\sum_{l=1}^{L} w_{hl}^{(2)} s_{zhl}^2 + \sum_{h=1}^{L} w_h^{(2)}\left[\left(\hat{\bar{Z}}_{h,\mathrm{prop}} - \hat{\bar{Z}}_{\mathrm{prop}}\right)^2 + \sum_{l=1}^{L} w_{hl}^{(2)}\left(\hat{\bar{Z}}_{hl} - \hat{\bar{Z}}_{h,\mathrm{prop}}\right)^2\right]\right)
\end{aligned}
\qquad (8)
$$

(Appendix B), with

$$
\hat{S}_{zh}^2 = \sum_{l=1}^{L}\frac{m_{hl}'}{m_h'} s_{zhl}^2 + \sum_{l=1}^{L}\frac{m_{hl}'}{m_h'}\left(\hat{\bar{Z}}_{hl} - \hat{\bar{Z}}_{h,\mathrm{prop}}\right)^2, \quad s_{zhl}^2 = \frac{1}{m_{hl}-1}\sum_{j=1}^{m_{hl}}\left(z_{hlj} - \hat{\bar{Z}}_{hl}\right)^2,
$$

and $w_h^{(2)}$ and $w_{hl}^{(2)}$ the estimates of relative strata and substrata sizes, respectively, on the second occasion.

If standard double sampling formulas (3) and (4) are used accordingly for the variance of estimator (6) instead of (7) and (8),

ignoring proportional stratified subsampling, the variance of (6) tends to be overestimated more or less due to the usually higher precision of stratified sampling with proportional allocation compared to simple random sampling (de Vries, 1986). Therefore, (3) and (4) could be used as simplified variance formulas leading to more or less conservative estimates of the sampling error.

In case of our example, many of the L strata will most likely contribute only a few transition units to a stratum h on the second occasion in practice. Usually, most of the transition units will originate from the next younger or, in case of the youngest deciduous or coniferous stratum, from the respective oldest stratum. For all strata l that do not contribute transition units to h, one obtains $m'_{hl} = 0$ as well as $m_{hl} = 0$, and the respective summands vanish in formulas (6) to (8).

If, for smaller m'_{hl}, one obtains $0 \leq m_{hl} \leq 1$, the variance estimator s^2_{hl} is not defined, and the m'_{hl} units of that substratum within stratum h should be merged with another substratum, as it is done with the collapsed stratum estimator (Wolter, 1985), or eventually simple random sampling might be assumed within all strata of the second occasion, both leading to conservative estimates as was reasoned above.

Since increasing age will be the predominant reason for first-phase units to change strata, it might be reasonable to use the constraints

$$v_1 \leq v_2 \leq v_3 \leq v_4 \,(\mathrm{DEC1}-4)\,\mathrm{and}\,v_5 \leq v_6 \leq v_7 \leq v_8(\mathrm{CON1}-4),$$

that means monotonously increasing sampling proportions with increasing age classes in the optimization process. In general, this would more often require to establish additional terrestrial sample plots in the second phase rather than giving up existing ones, an exception being the transition of first-phase units from the two oldest age classes to the youngest, from stratum 4 to stratum 1 and from stratum 8 to stratum 5 after harvesting the plot.

5. Estimation of net change

The estimators $\hat{\bar{Y}}$ and $\hat{\bar{Z}}_{\mathrm{prop}}$ estimate the means on the first and second occasion, respectively. The estimator of net change is simply $\hat{\bar{Z}}_{\mathrm{prop}} - \hat{\bar{Y}}$, and for its variance it holds

$$V(\hat{\bar{Z}}_{\mathrm{prop}} - \hat{\bar{Y}}) = V(\hat{\bar{Z}}_{\mathrm{prop}}) + V(\hat{\bar{Y}}) - 2Cov(\hat{\bar{Y}}, \hat{\bar{Z}}_{\mathrm{prop}})$$

$$Cov(\hat{\bar{Y}}, \hat{\bar{Z}}_{\mathrm{prop}}) = \sum_{h=1}^{L}\sum_{k=1}^{L} \frac{n'_h}{n'} \frac{m'_k}{m'} \quad Cov(\hat{\bar{Y}}_h, \hat{\bar{Z}}_{k,\mathrm{prop}}) = \sum_{h=1}^{L}\sum_{k=1}^{L} \frac{n'_h}{n'} \frac{m'_k}{m'} \frac{1}{n_h} \frac{1}{m_k} \sum_{i=1}^{n_h}\sum_{j=1}^{m_k} Cov(y_{hi}, z_{kj}). \tag{9}$$

Herein, the covariance between $\hat{\bar{Y}}_h$ and $\hat{\bar{Z}}_{k,prop}$ can be estimated by

$$C\hat{o}v(\hat{\bar{Y}}_h, \hat{\bar{Z}}_{k,\mathrm{prop}}) = \frac{1}{n_h}\frac{1}{m_k} u_{hk} \frac{1}{u_{hk}-1} \sum_{r=1}^{u_{hk}} (y_{hr} - \bar{y}_{h(u_{hk})})(z_{kr} - \bar{z}_{k(u_{hk})}),$$

where u_{hk} is the number of second-phase plots which shifted from stratum h to stratum k (or stayed in h if $h = k$), and $\bar{y}_{h(u_{hk})}$ and $\bar{z}_{k(u_{hk})}$ are the means of those u_{hk} plots on occasions one and two, respectively. If $u_{hk} = 0$, the covariance is 0, if $u_{hk} = 1$, it cannot be estimated but may be set to 0, particularly if $n_h m_k$ is large.

The development of the true covariance formula in (9) is straight forward, and the estimator of the covariance between $\hat{\bar{Y}}_h$ and $\hat{\bar{Z}}_{h,\mathrm{prop}}$ comprises the factor $(n_h m_k)^{-1}$ of the true covariance and an estimator of the total of $n_h m_k$ terms $Cov(y_{hi}, z_{kj})$. Only u_{hk} of these terms are unequal 0 and can be estimated by

$$\frac{1}{u_{hk}-1} \sum_{r=1}^{u_{hk}} (y_{hr} - \bar{y}_{h(u_{hk})})(z_{kr} - \bar{z}_{k(u_{hk})})$$

This finally yields the proposed covariance estimator.

6. Optimization

Usually, optimum allocation of sample plots to strata is achieved by minimizing the inventory costs for a predefined variance of $\hat{\bar{Y}}$, or by minimizing that variance for given total costs (Cochran, 1977). Instead of fixing the costs, we minimize the variance for a fixed terrestrial sample size n. We assume equal costs per unit in all strata, because the inventory teams are often paid per plot, independent of the plot type or the amount of time spent at the plot. Therefore, n is an appropriate placeholder for the inventory costs. The optimum allocation is then estimated by

$$v_h = \frac{n_h}{n'_h} = n \frac{s_h}{\displaystyle\sum_{h=1}^{L} n'_h s_h}$$

(Appendix C). An optimum allocation can be obtained using the inventory data of the first occasion for the calculation of s^2_h. It will be an appropriate guideline for planning the repeated inventory on the second occasion if the within-strata variances do not change remarkably during one inventory cycle, which is about 10 years in forest district inventories in Lower Saxony.

Optimization needs to focus on target variables and populations of primary interest, and it usually has to serve multiple purposes in practice. Since target diameter selection has become a wide-spread harvesting strategy, foresters are often particularly interested in more precise estimations of volume per hectare of higher diameter classes. Discussions with forest management officers lead to the definition of 4 target populations: Oak, dbh > 50 cm; Beech, dbh > 50 cm; Spruce, dbh > 35 cm, and Pine, dbh > 40 cm. They represent the most important tree species groups in the example district (Oak 16.32 m^3/ha, Beech 86.57 m^3/ha, Spruce: 38.66 m^3/ha, Pine: 0.74 m^3/ha) and the operationally most relevant diameter classes. Allocation of sampling units was to be optimized with respect to these populations. According to the recommendation of Cochran (1977, 5A.3) and de Vries (1986, 2.4), optimal v_h were calculated for each of the populations and finally averaged (Table 2b).

As a reference for the discussion of sampling errors under different allocations of second-phase sample plots, we calculated also the sampling errors of a simple random sampling design according to

$$V(\hat{\bar{Y}}) = \frac{S^2}{n} \tag{10}$$

using the variance decomposition (Appendix A)

$$S^2 = \frac{1}{\lambda(F)} \sum_{h=1}^{L} \int_{F_h} \left(Y(x) - \bar{Y}_h + \bar{Y}_h - \bar{Y}\right)^2 dx$$

$$= \sum_{h=1}^{L} \frac{\lambda(F_h)}{\lambda(F)} S^2_h + \sum_{h=1}^{L} \frac{\lambda(F_h)}{\lambda(F)} \left(\bar{Y}_h - \bar{Y}\right)^2$$

The strata weights were estimated by $w_h = n'_h/n'$, the within-strata variances by s^2_h, and we substituted $\hat{\bar{Y}}_h$ and $\hat{\bar{Y}}$ for the according population parameters. This was equivalently done for 2008, where we used the stratification of 1999 and the sample plot data of 2008.

Table 4
Final number m_{hl} of second-phase plots using the 1999 allocation in 2008, number of plots to be established in 2008 (+), and number of existing plots to be released (−).

		Stratum l in 1999							
		DEC1	DEC2	DEC3	DEC4	CON1	CON2	CON3	CON4
Stratum h in 2008	DEC1	98 (−6)	3 (+1)	1	10	4	3 (−3)	5 (−6)	1 (−1)
	DEC2	30 (−11)	99 (+1)	1 (−1)	3 (−2)	4 (−6)	8 (−22)	1 (−4)	0
	DEC3	1	32 (+9)	110 (−8)	6 (−3)	1 (−2)	4 (−12)	4 (−6)	0 (−1)
	DEC4	14 (+1)	6 (+3)	41 (+21)	326 (−2)	1 (+1)	2 (+1)	3 (−1)	1
	CON1	6 (+5)	1 (−1)	0	1 (+1)	38 (−1)	1 (+1)	3 (+2)	1
	CON2	6 (+4)	16 (+10)	4 (+2)	5 (+2)	128 (+40)	505 (−9)	3	2 (−1)
	CON3	3 (+3)	1 (+1)	1 (+1)	1	3 (+2)	37	44 (+2)	0
	CON4	0	1	0	3 (+3)	1	0 (−1)	10 (+1)	3 (−1)

Table 5
Final number m_{hl} of second-phase plots using the optimized allocation of 1999 in 2008, number of plots to be established in 2008 (+), and number of existing plots to be released (−).

		Stratum l in 1999							
		DEC1	DEC2	DEC3	DEC4	CON1	CON2	CON3	CON4
Stratum h in 2008	DEC1	58 (−46)	2	1	5 (−5)	2 (−2)	1 (−5)	2 (−9)	1 (−1)
	DEC2	57 (+16)	189 (+91)	3 (+1)	6 (+1)	8 (−2)	15 (−15)	1 (−4)	0
	DEC3	2 (+1)	48 (+25)	164 (+46)	8 (−1)	1 (−2)	7 (−9)	6 (−4)	0 (−1)
	DEC4	23 (+10)	10 (+7)	66 (+46)	529 (+201)	1 (+1)	4 (+3)	4	2 (+1)
	CON1	3 (+2)	1 (−1)	0	0	15 (−24)	0	0 (−1)	0 (−1)
	CON2	2	6	1 (−1)	2 (−1)	49 (−39)	192 (−322)	1 (−2)	1 (−2)
	CON3	4 (+4)	1 (+1)	0	2 (+1)	4 (+3)	47 (+10)	54 (+12)	0
	CON4	1 (+1)	0 (−1)	0	4 (+4)	2 (+1)	1	16 (+7)	6 (+2)

7. Case study

The sampling proportions optimized using the 1999 inventory data are provided in Table 2a. Because of the usually larger variances in older strata, it could be expected beforehand that the optimum allocations tend to be higher in older strata. This is clearly confirmed by the result in Table 2a, except for the two intermediate age classes of the deciduous species strata, where the monotony constraint lead to 0.13 for both strata instead of the original optima 0.15 and 0.11 for DEC2 and DEC3, respectively.

If the allocation ν_h of second-phase units according to Table 2a (1999) was also used in 2008 for μ_h, 111 of 1644 existing second-phase plots would have been released in 2008, and 118 new plots established (Table 4). With the μ_h optimized using the 1999 data, 501 plots and 498 plots would have been released and newly established, respectively (Table 5).

The resulting sampling errors for the two allocations, original and optimized, as well as for a simple random sampling scheme which uses the same number of terrestrial (second-phase) plots, are reported in Table 6a and b. We used the estimator (5) and variances (7) and (10). In (7) and (10), we substituted estimates for means and variances based on the existing second-phase plots established 1999. That means $\bar{Z}_h$ is not estimated by $\hat{\bar{Z}}_{h,\text{prop}} = 1/m_h \sum_{l=1}^{L} \sum_{j=1}^{m_{hl}} z_{hlj}$, because using only the second-phase plots of 1999 we have a non-proportional allocation to the substrata and must apply an appropriate weighting for the substrata sample means. Instead, we use $\hat{\bar{Z}}_h = 1/m_h \sum_{l=1}^{L} \left(m'_{hl}/m'_h \right) (1/m_{hl}) \sum_{j=1}^{m_{hl}} z_{hlj}$. Further, for the variances of 2008 as given in Table 6b, we merged substrata with $m_h = 0$ or $m_h = 1$ with the next younger or next older substratum into a larger one allowing for the calculation of within-substrata variances s_{zhl}^2. All three standard errors of each population were related to a unique volume per hectare. It was estimated by (2) and $\hat{\bar{Z}} = \sum_{h=1}^{L} \left(m'_h/m' \right) \hat{\bar{Z}}_h$ ($\hat{\bar{Z}}_h$ as previously defined) for the first and second occasion, respectively.

Differences between the sample copy variance estimator (4) and the unbiased estimator (5) for the relative variances in Table 6a were small ($\leq 10^{-4}$, fourth digit or smaller) for all 24 estimates.

Mean ranks of the three designs (1999 and optimized allocation, simple random sampling) in 1999 for the 4 target populations are 9/4, 5/4 and 10/4, respectively. In 2008 they are still nearly unchanged (8/4, 5/4 and 11/4), with the optimized allocation having the smallest mean rank. This is also true for the 8 highest diameter classes with mean ranks 16/8, 13/8 and 19/8 in 2008. Over all diameter classes, mean ranks 24/12, 24/12 and 24/12 are all equal in 2008.

To assess the efforts necessary to achieve the precision of the optimized 2SS design in 2008 using a simple random sampling

Table 6
Comparison of relative sampling errors (ranks per population) of 1999 allocation, optimal allocation and simple random sampling in 1999 (a) and 2008 (b). Target populations of the optimization shaded.

Target population	1999 allocation	Optimized allocation	Simple random sampling
(a)			
Beech < 25 cm	0.0613 (3)	0.0588 (2)	0.0522 (1)
Beech 25–50 cm	0.0405 (3)	0.0328 (1)	0.0377 (2)
Beech > 50 cm	0.0389 (2)	0.0353 (1)	0.0429 (3)
Oak < 25 cm	0.1728 (2)	0.1755 (3)	0.1366 (1)
Oak 25–50 cm	0.1639 (3)	0.1247 (2)	0.1219 (1)
Oak > 50 cm	0.1273 (3)	0.1009 (1)	0.1149 (2)
Spruce < 25 cm	0.0581 (1)	0.0814 (2)	0.0747 (1)
Spruce 25–35 cm	0.0399 (1)	0.0540 (2)	0.0542 (3)
Spruce > 35 cm	0.0420 (1)	0.0479 (2)	0.0583 (3)
Pine < 25 cm	0.2422 (1)	0.3895 (3)	0.3296 (2)
Pine 25–40 cm	0.1862 (1)	0.2481 (3)	0.2157 (2)
Pine > 40 cm	0.3734 (3)	0.2983 (1)	0.3394 (2)
(b)			
Beech < 25 cm	0.0576 (3)	0.0563 (2)	0.0512 (1)
Beech 25–50 cm	0.0380 (3)	0.0310 (1)	0.0356 (2)
Beech > 50 cm	0.0396 (2)	0.0344 (1)	0.0414 (3)
Oak < 25 cm	0.1843 (2)	0.2226 (3)	0.1643 (1)
Oak 25–50 cm	0.1389 (3)	0.1109 (2)	0.1065 (1)
Oak > 50 cm	0.1170 (3)	0.0925 (1)	0.1060 (2)
Spruce < 25 cm	0.0815 (1)	0.1139 (3)	0.1035 (1)
Spruce 25–35 cm	0.0445 (1)	0.0600 (2)	0.0621 (3)
Spruce > 35 cm	0.0371 (1)	0.0419 (2)	0.0495 (3)
Pine < 25 cm	0.4179 (2)	0.4441 (3)	0.3952 (1)
Pine 25–40 cm	0.1889 (1)	0.2440 (3)	0.2238 (2)
Pine > 40 cm	0.2636 (2)	0.2514 (1)	0.2796 (3)

Table 7
Comparison of double sampling for stratification and simple random sampling. Total (n) and additional (n^+) sample sizes to achieve the precision of 2SS ($n' = 6803$, $n = 1644$, optimum allocation) with simple random sampling for the 4 target populations in 2008.

Population	n	n^+
Oak, dbh > 50 cm	2,156	512
Beech, dbh > 50 cm	2,377	733
Spruce, dbh > 35 cm	2,289	645
Pine, dbh > 40 cm	2,130	486

design, we calculated the number of sample plots to be established and measured in the random sampling design (Table 7). They have to be compared with $n = 1644$ and the additional costs for the 2SS. The latter must be assessed within the wider framework of the inventory. If, as practiced in Lower Saxony, the CIR images are also used for other purposes such as delineation of management units, calamity assessment, and spatial prediction of stand parameters, related costs should not or at least not fully be considered in an efficiency analysis. Currently, 35 Euro are paid per second-phase plot and 0.35 Euro per first-phase unit (excluding acquisition of CIR images).

Finally, net volume change is estimated as -7.64 m^3/ha, based on the 1999 plots, and the estimated sampling errors are 3.83 m^3/ha for the 1999 allocation and 3.24 m^3/ha for the optimized allocation.

8. Discussion

Under the infinite population approach for the 2SS design, the sample of first-phase sample points is randomly taken from the infinite number of points located in the study area, whereas the second-phase sampling is conditionally done from finite populations, namely the n'_h first-phase sample points assigned to stratum h. The variance expression of the 2SS estimate of the population mean is simpler than the variance in the finite population approach, but it approximately equals the latter for large finite population size N and relatively small first-phase sample size n'. The variance estimator in the infinite population approach is unbiased and differs only slightly from the sample copy of the true variance, the difference being negligible for moderately large n'_h. In the case study, differences between approximate and unbiased estimates of the relative sampling errors were $\leq 10^{-4}$ for all target populations, in accordance with Cochran's (1977) statement.

On the second occasion, the 2SS is expanded by an additional subsampling with proportional allocation. Again we could derive an unbiased estimator for the population mean and its true variance, and we propose the sample copy of the variance formula as a variance estimator which will not be exactly unbiased but consistent. It is justified by Cochran's recommendation of the sample copy estimator in the simple 2SS design and our own empirical result for the first occasion.

The advantage of 2SS over simple random sampling regarding the sampling error is not only due to stratification but also to an appropriate allocation of second-phase units to the strata depending on variances within and between strata. If inventory costs per unit are homogeneous, the latter will be a non-proportional allocation, whenever within-strata variances are inhomogeneous (Cochran, 1977). In our case study, we suggested an unequal allocation to the 8 strata which was the mean of optimal allocations for 4 target populations. The sampling proportions suggested by that ad-hoc method increased with increasing age class (Table 2a), except for the strata DEC2 and DEC3, where the optimization resulted in a slight decrease which was corrected afterwards. The optimized allocation turned out to achieve a remarkably higher precision compared to simple random sampling for the 4 target populations on both occasions in 1999 and 2008. Herewith, the results for 2008

are of higher practical relevance, because they show that the optimization based on the 1999 inventory data is still clearly superior to simple random sampling in 2008, regarding the target populations.

Table 6 shows also the trade-off between that higher precision for the target populations and lower precision for lower diameter classes. In 2008 the intermediate diameter classes of beech and spruce are still estimated more precisely by the optimized 2SS than with simple random sampling, but for the intermediate oak and pine diameter class, and all lowest diameter classes, simple random sampling is superior.

The 2SS of 2008 using the allocation applied in 1999 is superior to simple random sampling for the biggest beeches and pines, the intermediate pines, and all diameter classes of spruce, but inferior for all oak diameter classes, the intermediate and lowest beech classes, and the lowest pine class. Compared to the optimized 2SS it is superior for the spruces and lower pine classes, but inferior for the biggest and intermediate beech and oaks, and the smallest beech as well as the biggest pines.

The optimized 2SS is clearly the most favorable design if highest precision is to be achieved for the highest and intermediate diameter classes of the three dominating tree species if one considers relative sampling errors and mean ranks. Even over all diameter classes the optimized version achieves the same mean rank as the two competing designs. The 1999 allocation is the second most favorable regarding the highest diameter, as well as regarding the highest and intermediate classes, according to the mean ranks.

According to Table 7, 733 additional second-phase plots would be necessary for a simple random sampling design to be superior or at least equal to the optimized 2SS design. Still about 600 additional plots were necessary if the precision of 2SS is to be achieved approximately with simple random sampling at least for the target populations of spruce and oak, which are the second and third most important populations; pine > 40 cm is rather rare (0.74 m^3/ha). Since costs per second-phase plots are usually much higher than expenses for stratification of first-phase units by interpretation of CIR images, and CIR images can additionally be exploited for many other purposes, we state a higher efficiency of the optimized 2SS approach, compared to simple random sampling. Regarding the current costs per unit, the costs for the additional terrestrial plots (600*35 Euro = 21,000 Euro) are nearly 9 times the costs for first-phase stratification (6803*0.35 Euro = 2381 Euro).

Enhanced GIS-based digital stand records would further reduce first-phase costs in future, but may also reduce precision of the stratification, because stratification will then be based on stand and no longer on local plot information, making a difference particularly in mixed or uneven-aged stands.

A crucial point for inventory planning is how to determine the sampling proportions for the first and following occasions, because derivation of optimized sampling proportions requires sample plot data from the different strata in advance. On the first occasion one might use proportions ν_h which proved to be optimal in other, similar districts. As soon as inventory data are provided from the first occasion, calculation of individual optimum μ_h for the target district can be derived and suggested for the second occasion, assuming that they will still be "nearly" optimal on that occasion. This is what we did in the case study. If the initial allocation can be applied again on the second occasion (Table 4), only a negligible number of plots has to be skipped (111) or established (118) on the second occasion, compared to the total sample size of 1644. These numbers become remarkably large (501 skipped, 498 established) if the optimized allocation replaces that of 1999 (Table 5). More case studies are necessary to answer the question if a global optimum over several districts exists that can be expected to be a good choice for an initial allocation in other districts.

In forest inventory practice, $m' = n'$ as used in our case study will often not hold, particularly if forest land is sold or new forest land

acquired. All first- and second-phase units of sold land are released from the analysis, and new first-phase units must be established and stratified on acquired land. All those new units belonging to a stratum h form an additional substratum, where second-phase plots have to be selected according to the inclusion probability μ_h. The same procedure is applied with first-phase units considered as forest land on the second occasion but not stratified on the first occasion or vice versa, caused by erroneous interpretation on one of the two occasions.

On the third and later occasions, the subsampling procedure as proposed in this paper can be applied accordingly. First-phase plots in each stratum h of occasion t are again stratified according to their origin on occasion $t-1$ and second-phase plots released or

$$
\begin{aligned}
S^2 &= \frac{1}{\lambda(F)}\sum_{h=1}^{L}\int_{F_h}\left(Y(x)-\bar{Y}_h+\bar{Y}_h-\bar{Y}\right)^2 dx \\
&= \frac{1}{\lambda(F)}\sum_{h=1}^{L}\left(\int_{F_h}\left(Y(x)-\bar{Y}_h\right)^2 dx + \lambda(F_h)\left(\bar{Y}_h-\bar{Y}\right)^2 + 2\left(\bar{Y}_h-\bar{Y}\right)\int_{F_h}\left(Y(x)-\bar{Y}_h\right)dx\right). \\
&= \sum_{h=1}^{L}\frac{\lambda(F_h)}{\lambda(F)}S_h^2 + \sum_{h=1}^{L}\frac{\lambda(F_h)}{\lambda(F)}\left(\bar{Y}_h-\bar{Y}\right)^2
\end{aligned}
$$

established randomly in order to achieve the proposed sampling proportion μ_h.

Acknowledgement

We would like to thank Dirk Weddig of the NW-FVA for careful stratification of first-phase units.

Appendix A. Proof of $V(\hat{\bar{Y}})$ and $\hat{V}(\hat{\bar{Y}})$

The variance of $\hat{\bar{Y}}$ can be derived following the proof of Särndal et al. (2003) for the so-called π^*-estimator using the well-known decomposition

$$V(\hat{\bar{Y}}) = V_1(E_2\hat{\bar{Y}}) + E_1 V_2(\hat{\bar{Y}}).$$

The subscripts 1 and 2 denote averaging over all first-phase and all second-phase samples given the first-phase sample, respectively. For the first term we get

$$E_2\hat{\bar{Y}} = \sum_{h=1}^{L}w_h\hat{\bar{Y}}_h' = \frac{1}{n'}\sum_{x\in\sigma'}Y(x),$$

wherein the prime at $\hat{\bar{Y}}_h$ indicates that the mean of local densities $Y(x)$ is calculated over all n_h' second-phase units of stratum h, measured as well as unmeasured. That yields, according to the one-phase one-stage scheme in Mandallaz (2008),

$$V_1(E_2\hat{\bar{Y}}) = \frac{1}{n'}\frac{1}{\lambda(F)}\int_F\left(Y(x)-\bar{Y}\right)^2 dx =: \frac{1}{n'}S^2.$$

For the second term of the variance decomposition, Särndal et al. (2003, equation (9.4.11)) particularly proved that

$$E_1 V_2(\hat{\bar{Y}}) = E_1\sum_{h=1}^{L}w_h^2 V_2(\hat{\bar{Y}}_h) = E_1\left(\sum_{h=1}^{L}\frac{w_h^2 s_h'^2}{n_h}\left(1-\frac{n_h}{n_h'}\right)\right) = E_1\left(\sum_{h=1}^{L}\frac{w_h s_h'^2}{n'}\left(\frac{1}{\nu_h}-1\right)\right)$$

$$s_h'^2 = \frac{1}{n_h'-1}\sum_{x\in F_h\cap\sigma'}\left(Y(x)-\hat{\bar{Y}}_h'\right)^2$$

Thus, we have

$$V(\hat{\bar{Y}}) = \frac{1}{n'}S^2 + E_1\left(\sum_{h=1}^{L}\frac{w_h s_h'^2}{n'}\left(\frac{1}{\nu_h}-1\right)\right). \tag{A.1}$$

This result can also be expressed in the form

$$V(\hat{\bar{Y}}) = \frac{1}{n'}\left(\sum_{h=1}^{L}W_h S_h^2 + \sum_{h=1}^{L}W_h\left(\bar{Y}_h-\bar{Y}\right)^2 + E_1\left(\sum_{h=1}^{L}w_h s_h'^2\left(\frac{1}{\nu_h}-1\right)\right)\right). \tag{A.2}$$

because

Cochran (1977, Theorem 12.3) states, by conditioning on the w_h, that $Ew_h s_h^2 = W_h S_h^2$, and the same argument leads to

$$E_1\left(\sum_{h=1}^{L}\frac{w_h s_h'^2}{n'}\left(\frac{1}{\nu_h}-1\right)\right) = \sum_{h=1}^{L}\frac{W_h S_h^2}{n'}\left(\frac{1}{\nu_h}-1\right). \tag{A.3}$$

Substituting this expression in Eq. (A.2) finally yields

$$V(\hat{\bar{Y}}) = \frac{1}{n'}\left(\sum_{h=1}^{L}\frac{W_h S_h^2}{\nu_h} + \sum_{h=1}^{L}W_h\left(\bar{Y}_h-\bar{Y}\right)^2\right).$$

For the proof of $\hat{V}(\hat{\bar{Y}})$ we use (Cochran, 1977, (12.27))

$$\sum_{h=1}^{L}w_h\left(\hat{\bar{Y}}_h-\hat{\bar{Y}}\right)^2 = \sum_{h=1}^{L}w_h\hat{\bar{Y}}_h^2 - \hat{\bar{Y}}^2,$$

and by conditioning on the w_h ($\vec{w}=(w_1,\ldots,w_L)'$) we find here

$$
\begin{aligned}
E\left(\sum w_h\hat{\bar{Y}}_h^2\,\Big|\,\vec{w}\right) &= \sum w_h\left[V(\hat{\bar{Y}}_h\,|\,\vec{w}) + \left(E(\hat{\bar{Y}}_h\,|\,\vec{w})\right)^2\right] \\
&= \sum w_h\left[\frac{1}{n_h}S_h^2+\bar{Y}_h^2\right] = \sum\frac{S_h^2}{n'\nu_h} + \sum w_h\bar{Y}_h^2 \\
E\left(\sum w_h\hat{\bar{Y}}_h^2\right) &= \sum\frac{S_h^2}{n'\nu_h} + \sum W_h\bar{Y}_h^2.
\end{aligned}
$$

It holds $V(\hat{\bar{Y}}_h\,|\,\vec{w}) = S_h^2/n_h$, because with the w_h fixed and the first-phase sample still varying, the n_h second-phase plots are selected from the infinite population F_h, and the result coincides with Cochran's equation (12.29) for large N.

Substitution of that result in the preceding formula yields

$$E\sum_{h=1}^{L}w_h\left(\hat{\bar{Y}}_h-\hat{\bar{Y}}\right)^2 = \sum\frac{S_h^2}{n'\nu_h} + \sum W_h\bar{Y}_h^2 - V(\hat{\bar{Y}}) - \bar{Y}^2,$$

and we get

$$E\hat{V}(\hat{\bar{Y}}) = \frac{1}{n'-1}\left(\sum\frac{W_h S_h^2}{\nu_h} + \sum W_h \bar{Y}_h^2 - V(\hat{\bar{Y}}) - \bar{Y}^2\right)$$

$$= \frac{1}{n'-1}\left(n'V(\hat{\bar{Y}}) - V(\hat{\bar{Y}})\right) = V(\hat{\bar{Y}}).$$

Appendix B. Proof of $V(\hat{\bar{Z}}_{\mathbf{prop}})$ and $\hat{V}(\hat{\bar{Z}}_{\mathbf{prop}})$

As in Appendix A, we find

$$V(\hat{\bar{Z}}) = \frac{S_z^2}{m'} + E_1\sum_{h=1}^{L}\left(\frac{m'_h}{m'}\right)^2 V_2\left(\hat{\bar{Z}}_{h,\text{prop}}\right).$$

Substituting the conditional variance by the well-known variance formula of stratified random sampling with proportional allocation (Cochran, 1977), we obtain

$$E_1\sum_{h=1}^{L}\left(\frac{m'_h}{m'}\right)^2 V_2\left(\hat{\bar{Z}}_{h,\text{prop}}\right) = E_1\sum_{h=1}^{L}\left(\frac{m'_h}{m'}\right)^2\frac{1-m_h/m'_h}{m_h}\sum_{l=1}^{L}\frac{m'_{hl}}{m'_h}s'^2_{zhl}$$

$$= E_1\frac{1}{m'}\sum_{h=1}^{L}\frac{m'_h}{m'}\left(\frac{1}{\mu_h}-1\right)\sum_{l=1}^{L}\frac{m'_{hl}}{m'_h}s'^2_{zhl}$$

with

$$s'^2_{zhl} = \frac{1}{m'_{hl}}\sum_{j=1}^{m'_{hl}}\left(z_{hlj} - \bar{z}'_{hl}\right)^2.$$

Considering now the ANOVA decomposition already used in Appendix A

$$S_z^2 = \sum_{h=1}^{L}W_h S_{zh}^2 + \sum_{h=1}^{L}W_h\left(\bar{Z}_h - \bar{Z}\right)^2,$$

the variance of $\hat{\bar{Z}}_{\text{prop}}$ is

$$V(\hat{\bar{Z}}_{\text{prop}}) = \frac{1}{m'}\sum_{h=1}^{L}W_h S_{zh}^2 + \frac{1}{m'}\sum_{h=1}^{L}W_h\left(\bar{Z}_h - \bar{Z}\right)^2$$

$$+ E_1\frac{1}{m'}\sum_{h=1}^{L}\frac{m'_h}{m'}\left(\frac{1}{\mu_h}-1\right)\sum_{l=1}^{L}\frac{m'_{hl}}{m'_h}s'^2_{zhl}.$$

This yields immediately the sample copy variance estimator

$$\hat{V}(\hat{\bar{Z}}_{\text{prop}}) = \frac{1}{m'}\sum_{h=1}^{L}\frac{m'_h}{m'}\hat{S}_{zh}^2 + \frac{1}{m'}\sum_{h=1}^{L}\frac{m'_h}{m'}\left(\hat{\bar{Z}}_{h,\text{prop}} - \hat{\bar{Z}}_{\text{prop}}\right)^2$$

$$+ \frac{1}{m'}\sum_{h=1}^{L}\frac{m'_h}{m'}\left(\frac{1}{\mu_h}-1\right)\sum_{l=1}^{L}\frac{m'_{hl}}{m'_h}s^2_{zhl}.$$

Again, from the ANOVA decomposition follows accordingly

$$S_{zh}^2 = \sum_{l=1}^{L}W_{hl}S_{zhl}^2 + \sum_{l=1}^{L}W_{hl}\left(\bar{Z}_{hl} - \bar{Z}_h\right)^2,$$

what directly leads to the proposed estimator $\hat{S}_{zh}^2$ of S_{zh}^2.

Appendix C. Proof of optimum ν_h

In order to optimize the ν_h we use (A.1) and (A.3) yielding

$$V(\hat{\bar{Y}}) = \frac{S^2}{n'} + \sum_{h=1}^{L}\frac{W_h S_h^2}{n'}\left(\frac{1}{\nu_h}-1\right)$$

so that minimization of $V(\bar{y})$ for predefined n' is equivalent to minimization of $\sum W_h S_h^2/\nu_h$. If n is given in advance, we have for the costs C and the expected costs

$$C = c'_0 n' + \sum_{h=1}^{L}c_h n_h \quad En_h = n'\nu_h W_h$$

$$EC = c'_0 n' + n'\sum_{h=1}^{L}c_h\nu_h W_h.$$

According to the Cauchy-Schwarz inequality for the product of expected costs and $V(\hat{\bar{Y}})$ or equivalently for their essential compartments

$$\left(\sum_{h=1}^{L}c_h\nu_h W_h\right)\left(\sum_{h=1}^{L}\frac{W_h S_h^2}{\nu_h}\right) \geq \left(\sum_{h=1}^{L}W_h\sqrt{c_h}S_h\right)^2,$$

$V(\hat{\bar{Y}})$ is minimal if $\nu_h\sqrt{c_h}/S_h = \text{const}(h = 1, \ldots, L)$. It implies that

$$\frac{\nu_h\sqrt{c_h}}{S_h} = \frac{\sum_{h=1}^{L}c_h\nu_h W_h}{\sum_{h=1}^{L}c_h\nu_h W_h} \cdot \frac{\nu_h\sqrt{c_h}}{S_h} = \frac{\sum_{h=1}^{L}c_h(n_h/n'_h)W_h}{\sum_{h=1}^{L}\sqrt{c_h}W_h S_h}$$

and

$$\nu_h = \frac{S_h}{\sqrt{c_h}}\frac{\sum_{h=1}^{L}c_h(n_h/n'_h)W_h}{\sum_{h=1}^{L}\sqrt{c_h}W_h S_h}.$$

Using $c_h = c$ and estimating the relative strata sizes W_h by n'_h/n', and S_h by s_h, one finally obtains the proposed formula.

References

Bickford, C.A., Mayer, C.E., Ware, K.D., 1963. An efficient sampling design for forest inventory: the northeastern forest survey. Journal of Forestry 61, 826–833.
Cochran, W.G., 1977. Sampling Techniques. Wiley, New York.
de Vries, P.G., 1986. Sampling Theory for Forest Inventory. A Teach-Yourself Course. Springer, Berlin.
Frayer, W.E., 1979. Multi-level Sampling Designs for Resource Inventories. USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins.
Gabler, K., Schadauer, K., 2007. Ansätze und Stichprobenpläne nationaler Forstinventuren. Austrian Journal of Forest Science 124 (1), 105–133 (in German).
Gregoire, T., Valentine, H., 2008. Sampling Strategies for Natural Resources and the Environment. Chapman & Hall.
Mac Lean, C.D., 1972. Improving inventory volume estimates by double sampling on aerial photographs. Journal of Forestry 70, 748–749.
Mandallaz, D., 2008. Sampling Techniques for Forest Inventories. Chapman & Hall/CRC, Boca Raton.
Rao, J.N.K., 1973. On double sampling for stratification and analytical surveys. Biometrika 60 (1), 125–133.
Särndal, C.-E., Swensson, B., Wretman, J., 2003. Model Assisted Survey Sampling. Springer, Berlin.
Schreuder, H.T., Gregoire, T.G., Wood, G.B., 1993. Sampling Methods for Multiresource Forest Inventory. Wiley, New York.
Scott, C.T., Köhl, M., 1994. Sampling with partial replacement and stratification. Forest Science 40 (1), 30–46.

Singh, B.D., Singh, D., 1965a. Some remarks on double sampling for stratification. Biometrika 52 (3), 587–590.

Singh, D., Singh, B.D., 1965b. Double sampling for stratification on successive occasions. Journal of the American Statistical Association 60 (311), 784–792.

Thompson, S.K., 1992. Sampling. John Wiley and Sons, New York.

Ware, K.D., Cunia, T., 1962. Continuous forest inventory with partial replacement of samples. For. Sci. Monogr., 3.

Wolter, K.M., 1985. Introduction to Variance Estimation. Springer, New York.

Williams, M.S., 2001. Comparison of estimation techniques for a forest inventory in which double sampling for stratification is used. Forest Science 47 (4), 563–576.