About the benefits of poststratification in forest inventories

J. Saborowski[1], **J. Cancino**[2]

[1]*Faculty of Forest Sciences and Forest Ecology, Georg-August University, Göttingen, Germany*
[2]*Facultad de Ciencias Forestales, Universidad de Concepción, Concepción, Chile*

ABSTRACT: A large virtual population is created based on the GIS data base of a forest district and inventory data. It serves as a population where large scale inventories with systematic and simple random poststratified estimators can be simulated and the gains in precision studied. Despite their selfweighting property, systematic samples combined with poststratification can still be clearly more efficient than unstratified systematic samples, the gain in precision being close to that resulting from poststratified over simple random samples. The poststratified variance estimator for the conditional variance given the within strata sample sizes served as a satisfying estimator in the case of systematic sampling. The differences between conditional and unconditional variance were negligible for all sample sizes analyzed.

Keywords: poststratification; systematic sampling; simple random sampling; conditional variance

Poststratification is well known as a means of increasing the precision of estimates in unstratified sampling by incorporating additional information about strata weights in the final estimator. In general, stratification leads to more precise estimations than simple random sampling when relatively homogenous strata can be configured with large variability between strata. Poststratification involves assignment of units after selection of the sample. Compared to a priori stratification, the variance of the poststratification estimator is increased by the randomness of the sample size in each stratum.

If poststratification is combined with systematic sampling, the gain in precision can be suspected to be small when the spatial distribution of strata leads to a nearly proportional allocation of sampling units to the strata, because in that case systematic sampling is approximately self-weighting. Proportional allocation is often at least approximately achieved by spatial systematic sampling in forest inventories, even if the strata are hidden during sample selection.

Finally, the poststratification variance estimator might be a nearly unbiased estimator for the variance of estimates based on systematic poststratified sampling because appropriate stratification can remarkably reduce trends in the underlying spatial data.

Systematic sampling

The usual one-dimensional systematic sampling design divides the N units of the population in $k \geq 2$ clusters or classes $S_1,..., S_k$, where S_i comprises the units $i, i+k, i+2k,..., i+jk$ ($i+jk \leq N$), and then selects one of these S_i at random. Selecting so the first unit as 1 of k yields an unbiased estimate of the population mean when $N = n \times k$, but that estimate is biased when $N \neq n \times k$. The bias arises from the fact that some of the k systematic samples have sample size n and others sample size $n+1$. A variant of the method, circular systematic sampling, also called Lahiri's method, provides both a constant sample size and an unbiased sample mean (Bellhouse, Rao 1975; Cochran 1977), but destroys the systematic structure of the sample by combining units from two different clusters. According to Cochran (1977) the implications of those varying sample sizes in case $N \neq n \times k$ can be assumed negligible if n exceeds 50 and are unlikely to be relevant even when n is small.

In general, n can not be arbitrarily fixed in advance. If $N = n \times k + c$ ($c \geq 0$) and $c < k$, then there are c samples of size n+1 and k-c samples of size n. When $2k > c > k$, $c - k$ systematic samples have n+2 units and the remaining have n+1 units. In more extreme cases, the sample size finally obtained can over- or underride the desired one remarkably. For example, with N = 102 and n = 30 desired N / n = 102 / 30 = 3.4 is obtained and one can choose among k = 3 or k = 4 systematic samples. In the first case c = 12 and 3 samples of size n = 34 are obtained, in the second case (c = 2) two systematic samples of size 25 and two of size 26 exist.

In two dimensions, a natural extension of one-dimensional systematic sampling is sampling on a regular grid. Most frequently, square grids are used in practice, although triangular grids may often be superior (Cochran 1977; Matérn 1960). Here, variability of sample size is usually even greater than in the one-dimensional case. The different systematic samples may vary by much more than one unit in size. For example in a squared population with N = 102 × 102 = 10,404 units, drawing each tenth unit in both directions results in 100 different systematic samples of varying size, that is, 64 samples of size 100, 32 of size 110, and 4 of size 121. In sampling a nonrectangular area, variability of the sample size will further be increased depending on the irregularity of the particular shape of the area. With poststratification there is an additional variability of sample sizes within strata (Valliant 1993).

A well-known drawback of systematic sampling is the absence of an unbiased variance estimator. Thus, practitioners make use of the simple random sampling variance estimator or one of the alternatives offered in the literature (e.g. Wolter 1985). The simple random sampling variance estimator often overestimates the true variance because it does not consider the self-weighting property of systematic sampling in case of hidden strata or spatial trends. Then systematic sampling has similar properties as stratified sampling with proportional allocation of samples and poststratified variance estimators, might be less biased.

With simple random sampling and appropriately large population and sample sizes, the sample means can be expected to be approximately normally distributed. This does not hold for systematic sampling, where the number of possible samples decreases with increasing sample size (Madow, Madow 1944). Whereas with simple random sampling the variance of the sample mean monotonically decreases with increasing sample size, this is not true for systematic sampling. Instead, there is a decreasing trend with erratic fluctuation (Madow 1946).

Poststratification

Poststratification means assigning sampling units to strata after observation of the sample, i.e. stratification is imposed at the analysis stage rather than at the design stage (Stehman et al. 2003). Therefore, sample sizes within strata can not be fixed in advance but must be assumed random depending on the samples actually selected. This is an additional source of variation.

Poststratification is usually applied when additional information about strata sizes is available. In the ideal case this additional information comprises the true strata weights, which might be known from previous work or other external data sources (Cochran 1977; Smith 1991; Valliant 1993). As with *a priori* stratification, poststratification can be based on one or more classification variables defining the strata.

With large sample sizes and simple random sampling, and even more with systematic sampling, poststratification can be expected to correspond approximately to stratified sampling with proportional allocation. Usually, it is discussed as a method supposed to increase precision (Cochran 1977; Valliant 1993; Stehman et al. 2003), because it reduces selection biases by reweighting after sample selection (Smith 1991; Little 1993; Rao et al. 2002). Since systematic sampling might be expected to come closer to proportional allocation than simple random sampling, one might conjecture that the relative increase in precision by poststratification will be larger with simple random than with systematic sampling.

Ghosh and Vogt (1993) affirmed that the conditional variance, where the condition is a given sample allocation, is the proper instrument for comparing the poststratification mean with the regular simple random or systematic sampling mean as estimators of the true population mean. They observed that the poststratified mean is often superior to the regular mean when the conditional variance or the conditional mean square error is used for comparing both estimators (Ghosh, Vogt 1988). Holt and Smith (1979) affirmed that, in theory, neither the post stratification estimator nor the sample mean is uniformly best in all situations but empirical investigations indicate that post stratification offers protection against unfavourable sample configurations and should be viewed as a robust technique. As each stratum mean is weighted by the relative size of that stratum in the population, the post stratified estimator automatically corrects for any badly balanced sample.

Variances and variance estimation

The unconditional variance of the poststratified mean

$$\bar{y}_{st.post} = \sum_{h=1}^{L} W_h \, \bar{y}_h$$

with $\bar{y}_h$ the sample mean in stratum h and samples of size n randomly selected in a population with L strata is approximately

$$\sigma^2_{\bar{y}_{st.post.uncond}} \approx \frac{1}{n}\left(1 - \frac{n}{N}\right)\sum_{h=1}^{L} W_h S_h^2 + \frac{1}{n^2}\sum_{h=1}^{L}(1 - W_h)S_h^2 \quad (1)$$

where W_h and S_h^2 are, respectively, the relative size and the variance of stratum h (Cochran, 1977, 5A.42). The first term in equation (1) is the variance of the estimator $\bar{y}_{st}$ of the population mean in (pre)stratified random sampling with proportional allocation

$$\sigma^2_{\bar{y}_{st.prop}} = \frac{1}{n}\left(1 - \frac{n}{N}\right)\sum_{h=1}^{L} W_h S_h^2 \quad (2)$$

and the second represents the increase in variance that arises from the randomness of the n_h (Coch-ran 1977, p. 134 f.). It is evident that this term approximates zero when n→∞. Furthermore, if the S_h^2 do not differ greatly, the increase is about $(L - 1)/n$ times the variance for proportional allocation, ignoring the finite population correction. With $n >> L$ the increase due to the second term in equation (1) is small compared with equation (2).

Because of the randomness of the within strata sample sizes, the variance formulas for prestratified samples may be regarded as inappropriate (Williams 1962). However, although the variance of a poststratified estimator can be computed unconditionally (i.e., across all possible realizations of within strata sample sizes), inferences made conditionally on the achieved sample configuration are desirable (Valliant 1993). The conditional variance of the poststratified mean, that is the variance given the within strata sample sizes $n_1, ..., n_L$ is

$$\sigma^2_{\bar{y}_{st.post.cond}} = Var_{post}\left(\sum_{h=1}^{L} W_h \, \bar{y}_h \,|\, n_1, ..., n_L\right) =$$

$$= \sum_{h=1}^{L} \frac{W_h^2}{n_h} S_h^2 \left(1 - \frac{n_h}{N_h}\right) \quad (3)$$

The respective estimators of (1), (2) and (3) are obtained by simply substituting the estimator s_h^2 for S_h^2, e.g.

$$s^2_{\bar{y}_{st.post.cond}} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_h^2 \left(1 - \frac{n_h}{N_h}\right)$$

Instead of (1), Thompson (1992) presented an alternative approximation of the variance of the poststratified mean, namely

$$\frac{1}{n}\left(1 - \frac{n}{N}\right)\sum_{h=1}^{L} W_h S_h^2 + \frac{1}{n^2}\left(\frac{N - n}{N - 1}\right)\sum_{h=1}^{L}(1 - W_h)S_h^2$$

and he uses $s^2_{\bar{y}_{st.post.cond}}$ as the according variance estimator, which evidently estimates (only) the conditional variance given the sample allocation $n_1, ..., n_L$, what is but completely satisfactory because one is usually interested in the precision of an estimate based on the sample allocation actually obtained (Rao 1988).

With k systematic samples the i^{th} of which yields a simple mean $\bar{y}(i)$ and a poststratified mean $\bar{y}_{st.post}(i)$, the true variances of those estimators are by definition

$$\sigma^2_{\bar{y}_{sys}} = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}(i) - \frac{1}{k}\sum_{i=1}^{k}\bar{y}(i)\right)^2 \quad (4)$$

$$\sigma^2_{\bar{y}_{st.post.sys}} = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{st.post}(i) - \frac{1}{k}\sum_{i=1}^{k}\bar{y}_{st.post}(i)\right)^2 \quad (5)$$

Finally, the variance of the sample mean $\bar{y}$ in simple random sampling is denoted by

$$\sigma^2_{\bar{y}} = \frac{1}{n}\left(1 - \frac{n}{N}\right)\frac{1}{N - 1}\sum_{i=1}^{N}(y_i - \bar{y})^2 \quad (6)$$

and, based on k simple random samples, we use

$$\tilde{\sigma}^2_{\bar{y}} = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}(i) - \frac{1}{k}\sum_{i=1}^{k}\bar{y}(i)\right)^2 \quad (7)$$

$$\tilde{\sigma}^2_{\bar{y}_{st.post}} = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{st.post}(i) - \frac{1}{k}\sum_{i=1}^{k}\bar{y}_{st.post}(i)\right)^2 \quad (8)$$

for the simulated variances of simple and poststratified means. The ~ is used to symbolize the variances approximated by simulation; variances (4) and (5) are true variances because all k systematic samples are considered. In the simulation study equations (6) and (7) should give almost equal results.

Data base and virtual forest landscape

In order to carry out a large scale simulation study, it was intended to create an artificial population as close as possible to a real forest landscape. Therefore, volume data and actual forest coverage from a geographical information system of the Solling area (Lower Saxony, Germany) were used as the data base. Volume data stem from a forest district inven-

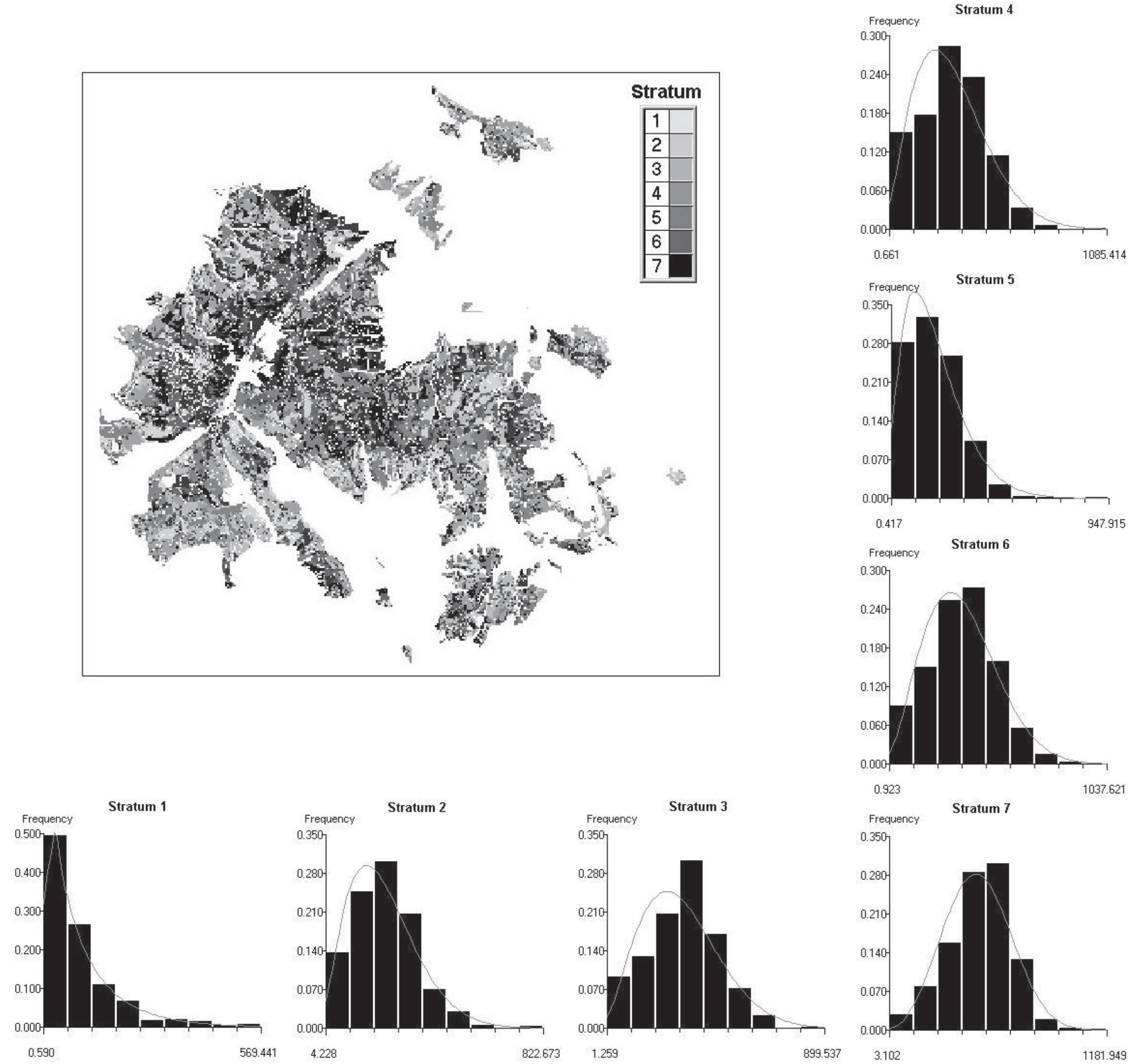

Fig. 1. Spatial coverage of the strata in the Solling, relative volume frequencies and fitted Weibull probability density function of each stratum

tory based on concentric circular plots where tree species and diameter in breast height of all sample trees are available as well as some heights required for calculating volumes (Böckmann et al. 1998). In total, data from 5,680 sample plots were incorporated in the creation of a virtual population.

The virtual population (Fig. 1) is represented by a mosaic of 212,386 squares (40*m* by 40*m* side length) each of which was assigned to one of 7 strata (Table 1) according to the stratum of the forest stand covering the centre of the square. Four strata were dominated by spruce (*Picea abies* [L.] Karst.) and three strata by beech (*Fagus sylvatica* L.).

Also, each inventory sample plot was assigned to one of the strata and a three-parameter Weibull function fitted to the volume per *ha* distribution of all sample plots of a stratum (Table 2). The Weibull parameters were estimated by the Maximum Likelihood method, with initial parameter values $\alpha = 0.95 \times V_{min}$, $\beta = V_{0.63} - \alpha$, and $\gamma = \beta/S_V$, where V_{min} is the minimum volume, $V_{0.63}$ represents the 63th percentile of volumes, and S_V is the standard deviation of the volume data. The resulting volume distributions range from negative exponential to left-skewed shapes (Fig. 1). From those volume distributions, the volume per ha for each square unit of the population was randomly selected depending on the stratum of the square unit. That implies in particular that trends, periodic variation or autocorrelation within strata are unlikely.

Simulation

Systematic samples were now chosen on square grids of 20 different grid widths representing sampling intensities from 0.047% to 1.0%. Those widths

Table 1. Characteristics of the 7 strata for Solling data

	Age class (years)	Stratum	N_h	W_h
Coniferous trees dominate	< 40	1	26,241	0.124
	41–80	2	30,801	0.145
	81–120	3	24,554	0.116
	> 120	4	39,002	0.184
Broadleaf trees dominate	< 40	5	30,690	0.145
	41–80	6	36,498	0.172
	> 80	7	24,600	0.116
		N =	212,386	

Table 2. Characteristic values of volume and estimated parameters of the three-parameter Weibull function per stratum

Stratum	Number of data points	Volume (m^3/ha)				Parameters of the Weibull function		
		Minimum	Maximum	μ	σ	α	β	γ
1	405	0.590	569.441	93.449	97.185	0.589938	89.195977	0.913904
2	372	4.228	822.673	225.129	119.641	4.228146	241.675473	1.764021
3	387	1.259	899.537	317.430	146.725	1.258811	343.767969	2.017091
4	894	0.661	1,085.414	314.517	165.170	0.661273	346.609165	1.831620
5	937	0.417	947.915	185.710	117.432	0.417108	201.749274	1.476572
6	1,658	0.923	1,037.621	348.028	160.364	0.922846	385.693802	2.159353
7	1,027	3.102	1,181.949	491.999	170.687	3.101742	535.959991	3.005017

were realized by selecting each 10th square in both directions for about 1% sampling intensity and each 46th square for 0.047%. Thus the number of systematic samples obtained varyed between $k = 100$ for the smallest and $k = 2{,}116$ for the largest grid width, sizes sufficiently large to obtain $n_h > 1$ in each stratum. For each of these intensities, the total number of different systematic samples were drawn, the values of the corresponding sampling units identified, and the simple ($\overline{y}$) and stratified ($\overline{y}_{st.post}$) means and the variance estimators for each sample as well as the true variances (4) and (5) calculated. Additionally, random samples (without replacement) of sample sizes equal to the mean sample sizes of the systematic samples were drawn and the corresponding $\overline{y}$, $\overline{y}_{st.post}$, the variance estimators as well as the "true" variances (7) and (8) calculated. All means and variances were averaged over the k systematic or random samples.

Sample sizes n vary among the k systematic samples and are constant among the k random samples. However, the within stratum sample sizes vary for both systematic and random sampling.

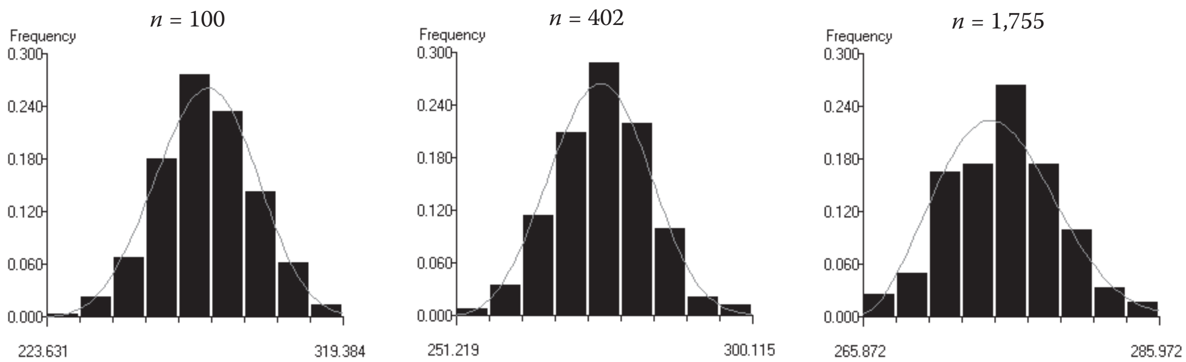

Fig. 2. Histogram of the poststratified sample mean $\overline{y}_{st.post}$ obtained from the corresponding k different systematic samples. Here n is the arithmetic mean of the sample size of the k samples in the population

Table 3. Characteristic values of systematic $\bar{y}_{st.post}$

Mean n	Number of systematic samples	Volume (m³/ha)			
		Minimum	Maximum	mean $\bar{y}_{st.post}$	$\sigma_{st.post.sys}$
100	2,116	223.631	319.384	275.394	15.300
147	1,444	239.977	314.349	275.431	12.161
195	1,089	238.964	318.364	275.487	11.333
252	841	245.358	304.173	275.479	9.523
291	729	252.316	302.744	275.428	8.828
340	625	254.066	302.227	275.417	8.143
402	529	251.219	300.115	275.457	7.665
439	484	255.799	298.070	275.441	7.419
482	441	256.247	292.334	275.382	6.593
531	400	258.862	294.995	275.477	6.395
588	361	259.578	293.504	275.407	5.889
656	324	260.168	293.516	275.448	6.081
735	289	259.209	288.234	275.459	5.607
830	256	262.688	290.079	275.450	5.103
944	225	262.346	287.911	275.435	4.796
1,084	196	263.943	288.117	275.466	4.314
1,257	169	266.995	284.735	275.462	4.049
1,475	144	265.965	284.841	275.447	3.976
1,755	121	265.872	285.972	275.433	3.670
2,124	100	267.151	284.702	275.424	3.004

RESULTS AND DISCUSSION

In theory, in a population with mean μ and variance σ^2, with simple random sampling without replacement and with large sample size, the distribution of the sample mean can be approximated by a normal distribution with mean μ and variance $(1 - n/N) \times \sigma^2/n$, independently of the original distribution of the variable of interest. Here, although the estimate of the true mean is unbiased and the variance of the mean decreases (Tables 3 and 4) with increasing n, its histogram approximates the normal probability density function (pdf) better for smaller than for the larger sample sizes (Fig. 2). This is due to the decreasing number k of systematic samples with increasing sample size n ($k = N/n$).

As expected (see chapter 2), the simulation confirmed the more or less erratic decrease of $\sigma_{\bar{y}\ st.post.sys}$ (Fig. 3a)

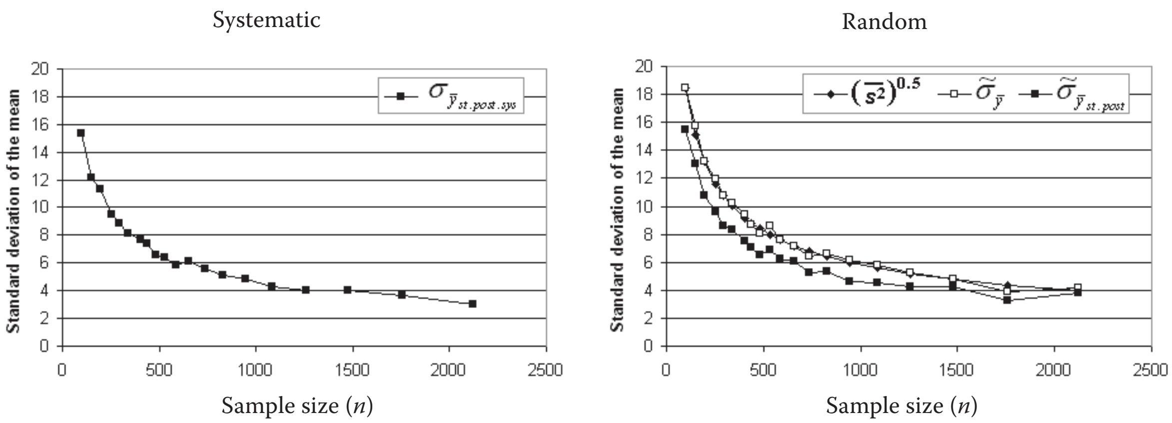

Fig. 3. Standard error of the poststratified mean for systematic and random sampling, the latter compared with the rooted mean variance estimate of the k replicated simple random samples and $\tilde{\sigma}_{\bar{y}}$ according to (7)

Table 4. Characteristic values of random $\bar{y}_{st.post}$

Sample size n	Number of random samples	Volume (m^3/ha)			
		Minimum	Maximum	mean $\bar{y}_{st.post}$	$\tilde{\sigma}_{st.post}$
100	2,116	221.263	339.764	275.452	15.505
147	1,444	237.849	315.398	275.122	13.060
195	1,089	239.759	311.636	274.785	10.768
252	841	246.214	307.115	275.199	9.615
291	729	249.701	299.559	274.906	8.587
340	625	251.328	303.244	275.899	8.288
402	529	250.967	299.248	274.377	7.504
439	484	256.562	292.695	275.097	7.067
482	441	259.611	294.242	275.809	6.560
531	400	257.124	296.154	275.342	6.834
588	361	258.846	292.583	275.493	6.279
656	324	256.534	292.863	275.233	6.090
735	289	259.901	293.623	275.352	5.243
830	256	258.429	290.645	275.150	5.318
944	225	262.949	287.922	275.372	4.644
1,084	196	261.287	292.002	275.442	4.547
1,257	169	264.998	287.275	275.283	4.270
1,475	144	264.997	287.718	275.699	4.237
1,755	121	266.267	284.233	275.063	3.294
2,124	100	267.991	284.271	275.344	3.761

with increasing sample size. The erratic behavior is more expressed for $n > k$, here beyond sample sizes of about 460, that is with sample sizes where $c > k$ might occur and where the variability of the sample size n decreases slower beyond that point (Fig. 4). Similar erratic oscillations of $\tilde{\sigma}_{\bar{y}_{st.post}}$ occur with random sampling, and the rooted mean variance estimate of the k replicated simple random samples and $\tilde{\sigma}_{\bar{y}}$ according to (7) exhibit no remarkable differences (Fig. 3b), although both are larger than $\tilde{\sigma}_{\bar{y}_{st.post}}$.

Fig. 5a compares the square root of the means of the estimates $s^2_{\bar{y}st.post.uncond}$ for the conditional variance (1), the means of $s^2_{\bar{y}\,st.post.cond}$ as estimates for the unconditional variance (2) and the means of the random sample variance estimates $s^2_{\bar{y}}$ with the true variance $\sigma^2_{\bar{y}\ st.post.sys}$ within the range of the analyzed sample sizes. Obviously, $s^2_{\bar{y}}$ overestimates the true variance by far, and the conditional and unconditional variance estimators, on an average, exhibit no remarkable differences. Thus, the component of variability associated to the variability of the sample

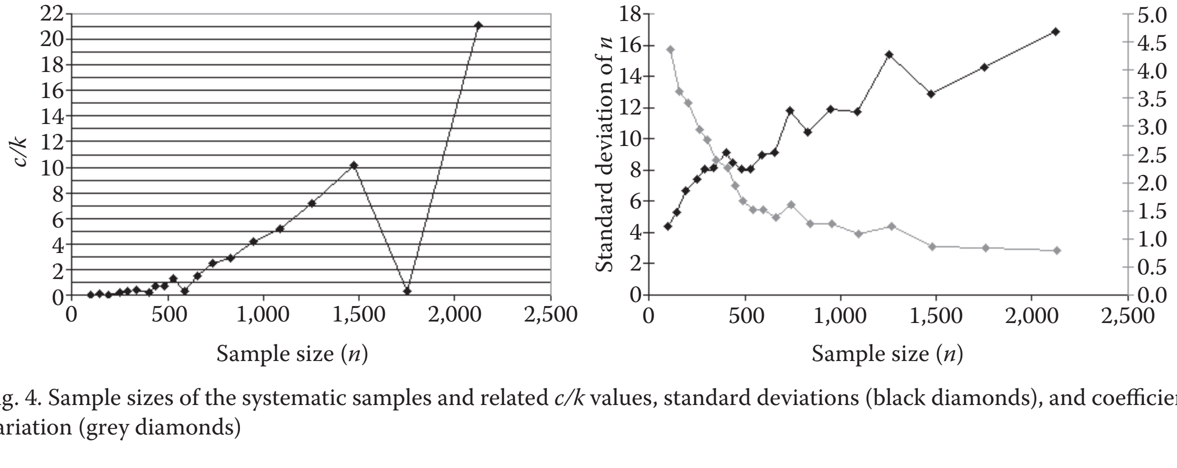

Fig. 4. Sample sizes of the systematic samples and related c/k values, standard deviations (black diamonds), and coefficients of variation (grey diamonds)

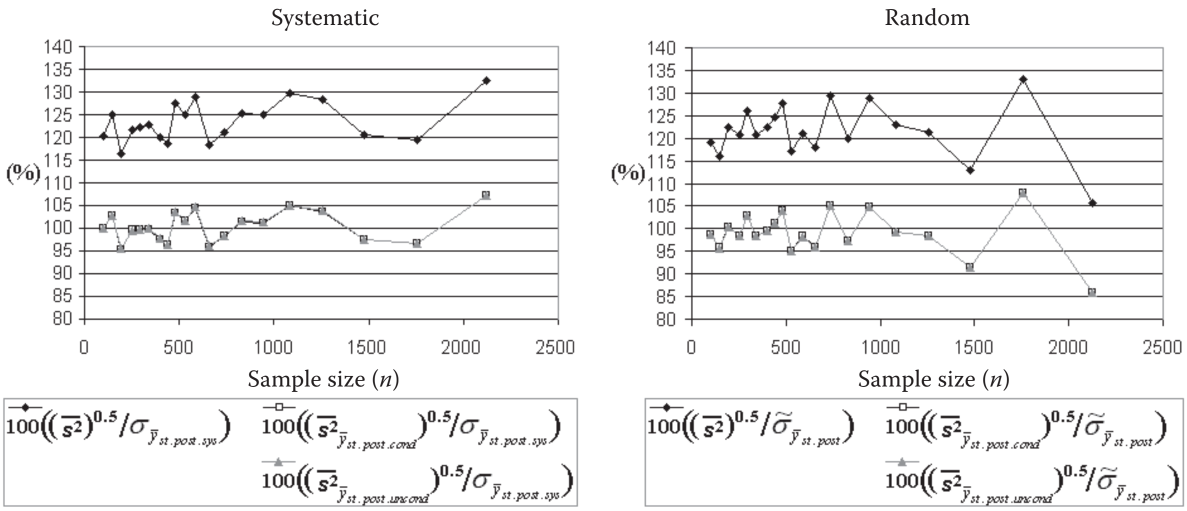

Fig. 5. 100+bias(%) of variance estimators for the true standard deviation of the poststratified mean in systematic and random sampling

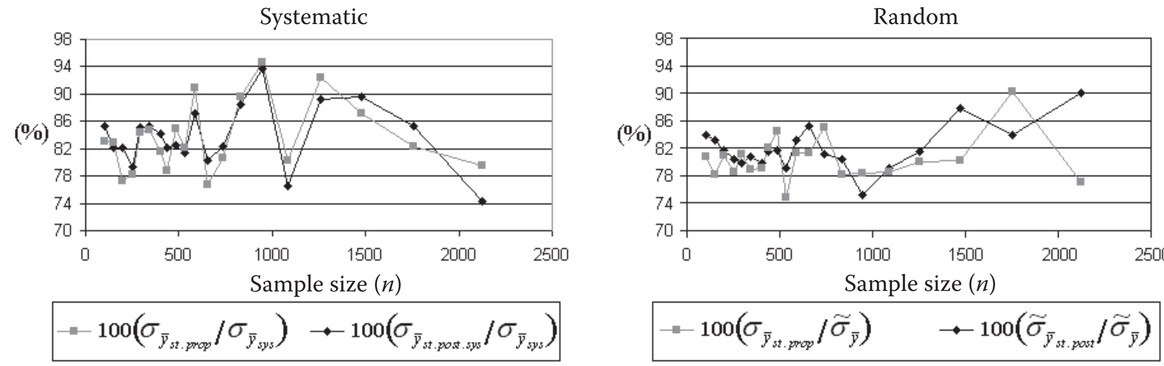

Fig. 6. Relative efficiency of poststratification in systematic and random sampling, real strata

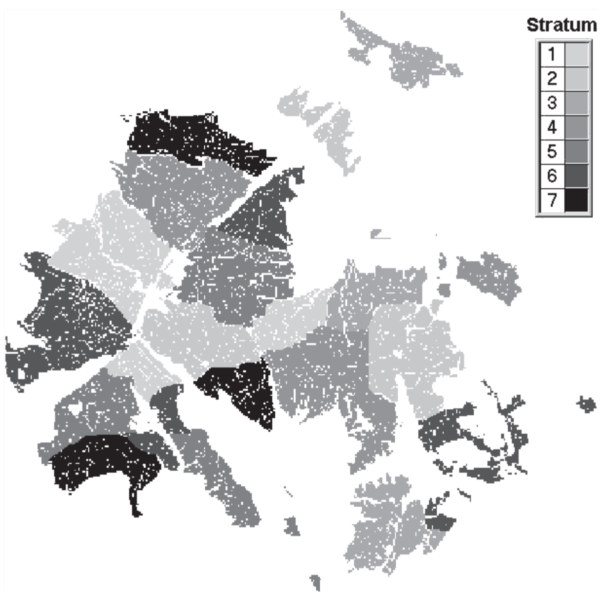

Fig. 7. Artificial strata with larger connected subareas

size is, as it was expected, practically zero. Biases are erratic, varying predominantly within a range of ± 5% of the true standard error of the systematic samples. Similar results can be observed with random sampling (Fig. 5b) where the same variance estimators are compared with the "true" variance $\sigma_{\bar{y}_{st.post}}$ of the poststratified mean.

Taking the true standard deviation $\sigma_{\bar{y}_{sys}}$ of the unstratified mean of a systematic sample as a reference, the standard deviation $\sigma_{\bar{y}_{st.post.sys}}$ of the poststratified mean under systematic sampling was about 16% smaller on the average (Fig. 6a). A similar gain in precision can be achieved by (pre)stratified sampling with proportional allocation in the underlying virtual forest landscape. Beyond sample sizes of about 500, that is of samples where *n* is larger than *k*, the variance ratios are less stable with gains in precision between 6 % and 25 %.

With random sampling (Fig. 6b), gains in precision are only slightly larger. Probably, the little size and spatial distribution of connected areas of the diffe-

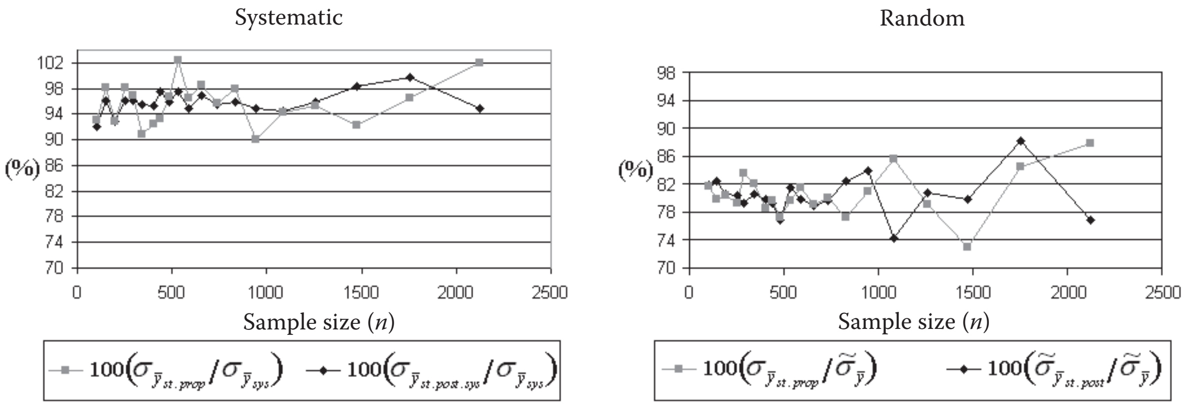

Fig. 8. Relative efficiency of poststratification in systematic and random sampling, artificial strata

rent strata leads to an allocation of the samples which is only a little closer to proportionality for systematic sampling than for random sampling. In that case reweighting by poststratification must have a similar effect for both sampling techniques.

In order to analyze the influence of the spatial structure of strata on the efficiency of poststratification, an artificial stratification was set up (Fig. 7). Here, the strata comprise larger connected subareas as for the real spatial distribution of strata (Fig. 1). The allocation of samples under systematic samples will be closer to proportionality in that case and should result in a lower relative efficiency of the poststratified mean (systematic sampling). This conjecture could be stated by the results presented in Fig. 8. Precision increased only by about 4%, instead of 16% before, for systematic sampling. For random sampling the increase of precision by poststratification remained at the same level as for the real stratification.

CONCLUSION

The case study presented reveals that mean estimators under systematic sampling can remarkably be improved in precision by poststratification when strata comprise a large number of small connected subareas. The larger connected subareas are the less is the gain in precision. The conditional as well as the unconditional variance estimator for poststratified sampling were only slightly biased (< 5%) with varying signs for different sample sizes, particularly in case of systematic random sampling. They can be expected practically identical in large scale forest inventories; here we studied sample sizes above 100.

For random sampling, the spatial structure of strata had no influence on the efficiency of poststratification compared to simple random sample means.

With the underlying population, stratified random sampling with proportional allocation and poststratified systematic sampling achieved similar precision, but this might be different when within strata variances vary more among strata than in this case study.

References

BELLHOUSE D.R., RAO J.N.K., 1975. Systematic sampling in the presence of a trend. Biometrika, *62*: 694–697.

BÖCKMANN T., SABOROWSKI J., DAHM S., NAGEL J., SPELLMANN H., 1998. A new conception for forest inventory in lower saxony. Forst und Holz, *53*: 219–226. (in German)

COCHRAN W.G., 1977. Sampling Techniques. New York, Wiley: 428.

GHOSH D., VOGT A., 1988. Sample configuration and conditional variance in poststratification. American Statistical Association Proceedings, Section on Survey Research Methods: 289–292.

GHOSH D., VOGT A., 1993. Some theorems relating poststratification and sample configuration. American Statistical Association Proceedings, Section on Survey Research Methods: 341–345.

HOLT D., SMITH T.M.F., 1979. Post stratification. Journal of the Royal Statistical Society A, *142*: 33–46.

LITTLE R.J.A., 1993. Post-stratification. A modeler's perspective. Journal of American Statistical Association, *88*: 1001–1012.

MADOW W.G, MADOW L.H., 1944. On the theory of systematic sampling. Annals of Mathematical Statistics, *15*: 1–24.

MADOW L.H., 1946. Systematic sampling and its relation to other sampling designs. Journal of American Statistical Association, *41*: 204–217.

MATÉRN B., 1960. Spatial variation. Meddelanden från statens Skogsforskningsinstitut, *49*: 1–144.

RAO J.N.K., 1988. Variance estimation in sample surveys. In: KRISHNAIAH P.R., RAO C.E. (eds.), Handbook of Statistics, Vol. 6 (Sampling). Amsterdam, Elsevier: 427–444.

RAO J.N.K., YUNG W., HIDIROGLOU M.A., 2002. Estimating equations for the analysis of survey data using poststratification information. Sankhya, *64* A: 364–378.

SMITH T.M.F., 1991. Post-stratification. The Statistician, *40*: 315–323.

STEHMAN S.V., SOHL T.L., LOVELAND T.R., 2003. Statistical sampling to characterize recent United States land-cover change. Remote Sensing of Environment, *86*: 517–529.

THOMPSON S.K., 1992. Sampling. New York, Wiley: 1–343.

VALLIANT R., 1993. Poststratification and conditional variance estimation. Journal of American Statistical Association, *88*: 89–96.

WILLIAMS W.H., 1962. The variance of an estimator with poststratified weighting. Journal of American Statistical Association, *57*: 622–627.

WOLTER K.M., 1985. Introduction to Variance Estimation. New York, Springer: 1–427.

Received for publication April 10, 2006
Accepted after corrections May 11, 2006

O přínosech poststratifikace v lesnické inventarizaci

ABSTRAKT: Na základě GIS databáze a údajů lesnické inventarizace pro určitý úsek lesa byl vytvořen rozsáhlý virtuální základní soubor. Tento soubor byl využit pro simulaci velkoplošné inventarizace s odhady parametrů získanými pomocí poststratifikace systematického a jednoduchého náhodného výběru a pro studium zvýšení přesnosti odhadu. Přes systematický výběr kombinovaný s poststratifikací se jeví stále ještě efektivnější než nestratifikovaný systematický výběr, zvýšení přesnosti se blíží výsledkům získaným z jednoduchého náhodného výběru s poststratifikací. Poststratifikovaný odhad rozptylu pro podmíněný rozptyl stanovený na základě velikosti výběrů jednotlivých oblastí (strat) slouží jako uspokojivý odhad v případě systematického výběru. Rozdíly mezi nepodmíněným a podmíněným rozptylem byly shledány pro všechny analyzované velikosti výběru jako zanedbatelné.

Klíčová slova: poststratifikace; systematický výběr; jednoduchý náhodný výběr; podmíněný rozptyl

Corresponding author:

Prof. Dr. Joachim Saborowski, Institut für Forstliche Biometrie und Informatik, Büsgenweg 4, 37077 Göttingen, Germany
tel.: + 49 551 393 450, fax: + 49 551 393 465, e-mail: jsaboro@gwdg.de