CrossMark

Spatially-balanced sampling versus unbalanced stratified sampling for assessing forest change: evidences in favour of spatial balance

Maria Chiara Pagliarella[1] · **Piermaria Corona[2]** · **Lorenzo Fattorini[1]**

Abstract Large-scale remote sensing-based inventories of forest cover are usually carried out by combining unsupervised classifications of satellite pixels into forest/non forest classes (map data) with subsequent time-consuming visual on-screen imagery classification of a probabilistic sample of pixels taken as the ground truth (reference data). In this paper the estimation of forest change from a sample of reference data is approached by: (i) exploiting map data to construct strata in which changes are occurred, and then adopting the stratified sampling joined with the HT estimator with most sampling effort devoted to strata where changes are occurred irrespective of their size, as suggested in most remote sensing literature regarding land change assessments; (ii) adopting a spatial scheme ensuring spatially balanced samples, as suggested in most recent statistical literature regarding spatial surveys, and exploiting the map data in the difference estimator. The results of a comparison performed on an artificial population of reference data generated from a real population of map data recorded in Sardinia (Italy) discourage the use of unbalanced stratified samples

✉ Maria Chiara Pagliarella
mariachiara.pagliarella@unisi.it

Piermaria Corona
piermaria.corona@crea.gov.it

Lorenzo Fattorini
lorenzo.fattorini@unisi.it

[1] Department of Economic and Statistics, University of Siena, Piazza San Francesco, 7, 53110 Siena, Italy

[2] Research Centre for Forestry and Wood, Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Viale Santa Margherita, 80, 52100 Arezzo, Italy

that achieve the worst precision. The best results are obtained by means of spatially balanced samples or stratification with nearly proportional allocation to strata.

Keywords Auxiliary information · Difference estimator · Forest monitoring · Horvitz–Thompson estimator · Simulation study · Spatially balanced sampling · Stratification

1 Introduction

Land cover changes are the expression of recurrent, consecutive land reorganizations in order to better adapt land-use and spatial structure to the varying changing environmental conditions and demands of human societies. Assessing forest cover change is one of the current main monitoring requirements of international environmental fora. For instance, *Reducing Emission from Deforestation and forest Degradation* (REDD) is a United Nations measure offering incentives to developing Countries for reducing emissions by increasing and improving forested lands: in this scenario, periodic and accurate estimates of forest changes are crucial (UN-REDD 2013).

Large-scale remote sensing-based inventories of forest cover are usually carried out by a combination of unsupervised classification of satellite imagery and subsequent manual (visual on-screen) imagery classification with the highest accuracy taken as ground truth (e.g. Hansen et al. 2013). Because visual on-screen operations are time expensive procedures, they can be carried out, as a rule, only for a probabilistic sample of pixels and forest cover is estimated from that sample. Under this protocol, the data arising from the unsupervised classification of the satellite pixels, henceforth referred to as *map data*, are available over the whole study region. Map data are likely to be good proxies of the on-screen classifications, henceforth referred to as *reference data*, which are instead available only for the sampled pixels (e.g. Sannier et al. 2014; Corona et al. 2015).

To accomplish the goal of efficiently estimating forest changes, any survey should be planned to provide reliable estimates of forest cover at any occasion of interest. Once an unbiased and efficient estimator of forest cover is adopted, the very natural way to estimate the change between two occasions is to take the difference of the forest cover estimates at the two occasions. In this case, the variance of the change estimator decreases as the covariance of the estimators at the two occasions increases. Thus, irrespective of the criterion adopted to estimate forest cover at the two occasions, a convenient way to plan remote-sensing based surveys is to maintain the same sample of pixels over time so that a positive covariance between the estimators of forest cover is induced. Many remote-sensing based surveys of forest cover as well as many land-cover and land-use surveys adopt this solution, which is usually referred to as the *pure panel*. Examples have been provided at global scale by the Global Forest Resources Assessment-FRA (http://www.fao.org/forestry/fra/remotesensing/en/), at continental scale by the Land Use/Cover Area frame statistical Survey in Europe-LUCAS (http://ec.europa.eu/eurostat/web/lucas/overview) and at Country scale by Sannier et al. (2014) and by the AGRIT project (http://www.itacon.it), among others.

A question is how to efficiently choose the sample of pixels where on-screen classification will be carried out. The question moves in the wide framework of spatial sampling. When sampling spatial units, the achievement of a so called *spatially balanced sample* (SBS), i.e. a sample in which units are well spread throughout the survey region, is the main target in most cases advised by statistical community. SBSs are usually constructed by avoiding or reducing the selection of contiguous units. There is a huge statistical literature on the ways of achieving SBS and some of them have appeared recently. We don't dwell here on these proposals because we'll come back to them later in the paper. On the other hand, in the remote sensing literature regarding land cover change assessment, completely reversed guidelines have been repeatedly provided, suggesting to unbalance the sample by a relative intensification of the sampling effort in the—usually small—strata where changes more likely occur (e.g. Richards et al. 2000; Baldauf et al. 2009; Stehman 2013; Olofsson et al. 2014). The purpose of this paper is to provide insights for choosing between these two opposite alternatives.

The paper is organized as follows. In Sect. 2 preliminaries and notations are given to state the problem. Map and reference data at the two survey occasions are defined and their different roles in estimation are outlined. The forest changes for both map and reference data are rewritten as the population means of the changes occurred within pixels. That allows to estimate the change directly from the differences recorded in the sampled pixels instead as the difference between the forest cover estimates at the two survey occasions (that is theoretically equivalent but more cumbersome to treat owing to the presence of the covariance between the estimators at the two occasions). In Sect. 3 the sampling strategy suggested by Olofsson et al. (2014) is considered. Stratification is carried out on the basis of changes occurred in the map data and samples of pixels are selected from each stratum by means of *simple random sampling without replacement* (SRSWOR). Then the *Horvitz–Thompson* (HT) criterion is adopted to estimate forest change. In Sect. 4, a strategy suggesting the achievement of SBSs is considered, irrespective of the changes occurred in map data. Among the several schemes providing SBSs, the so-called *one-per stratum-stratified sampling* (OPSS) is adopted owing to its simplicity and efficiency. Then the map data information is exploited at estimation level adopting the *difference* (D) criterion to estimate forest change. In Sect. 5 the two strategies are compared on the basis of the map data recorded in a study area of 400 km^2 settled in Sardinia (Italy) from which reference data are artificially generated (high-change population). Results of the simulation are reported in Sect. 6, while discussion and conclusions are reported in Sect. 7. In this last section reference is also made to results from a low-change population settled in Central Italy, whose details are omitted from brevity, that lead to conclusions very similar to the ones from the Sardinia test case.

2 Statement of the problem

Consider a survey area A covered by a grid of N satellite pixels. Denote by U the population of the N pixels. The pixel-level information about forest change from occasion 1 to occasion 2 is constituted by four dichotomous variables $x_{1j}, x_{2j}, y_{1j}, y_{2j}$.

The variables x_{1j} and x_{2j} arise from forest/non-forest satellite classification performed at both survey occasions: x_{tj} is equal to 1 if the j-th pixel at time t is classified as forest and is equal to 0 otherwise ($t = 1, 2$). Similarly the variables y_{1j} and y_{2j} arise from forest/non forest on-screen interpretation: y_{tj} is equal to 1 if the j-th pixel at time t is interpreted as forest and is equal to 0 otherwise ($t = 1, 2$). The x_{1j} s and x_{2j} s arise from satellite spectral classes regrouped into forest and non-forest thematic classes, and are readily available from satellite maps for all the pixels in the population. On the other hand, the y_{1j} s and y_{2j} s arise from time consuming forest expert works, based on satellite imagery in combination with available very high resolution imagery and Google Earth, and map archives like Bing Maps, Google Maps, national maps, local maps, etc. (e.g. Sannier et al. 2014). For these reasons, the x_{1j} s and x_{2j} s are referred to as the map data and can be used as auxiliary information in the estimation of forest cover, while the y_{1j} s and y_{2j} s are higher quality data, referred to as reference data, and will be taken as ground truth, stated that field data cannot be collected owing to the difficulty of access in dense forests. Owing to the high cost of reference data collection, these data cannot be known for all the pixels in the population but only for a sample of them.

Because all the map data are known at both survey occasions for each $j \in U$, their population means

$$\bar{X}_1 = \frac{1}{N} \sum_{j \in U} x_{1j} \text{ and } \bar{X}_2 = \frac{1}{N} \sum_{j \in U} x_{2j}$$

are also known and represent the fraction of grid area classified as forest from the satellite map at both survey occasions. Therefore their difference $D_X = \bar{X}_2 - \bar{X}_1$ is known and represents the increase of the fraction classified as forest between the two times. The difference D_X can be rewritten directly as

$$D_X = \frac{1}{N} \sum_{j \in U} d_{Xj}$$

where $d_{Xj} = x_{2j} - x_{1j}$ is the change occurred in the map data at pixel j.

Because the reference data cannot be known for each $j \in U$ owing to the recording costs, their population means

$$\bar{Y}_1 = \frac{1}{N} \sum_{j \in U} y_{1j} \text{ and } \bar{Y}_2 = \frac{1}{N} \sum_{j \in U} y_{2j}$$

are unknown and represent the fraction of grid area interpreted as forest from experts. Therefore, their difference $D_Y = \bar{Y}_2 - \bar{Y}_1$ is unknown and represents the increase of the fraction classified as forest between the two survey occasions. The difference D_Y can be rewritten directly as

$$D_Y = \frac{1}{N} \sum_{j \in U} d_{Yj}$$

where $d_{Yj} = y_{2j} - y_{1j}$ is the change occurred in the reference data at pixel j. The difference D_Y is taken as the *forest change* and constitutes the target parameter to be estimated from a sampling strategy.

On the basis of empirical investigations, Sannier et al. (2014) provide evidence that satellite classifications closely guess expert interpretations. The close matching between map and reference data suggests the use of d_{Xj} s as accurate and effective proxies for the d_{Yj} s to be exploited as auxiliary information.

3 Change estimation from stratified sampling

Map data information can be exploited at design level to construct strata on the basis of the changes recorded in the satellite map between the two survey occasions. Therefore, the population of N pixels is partitioned into four strata $U_{FF}, U_{FN}, U_{NF}, U_{NN}$ of size $N_{FF}, N_{NF}, N_{FN}, N_{NN}$, respectively, in accordance to the four possible combinations of values attained by the x_{1j} s and x_{2j} s.. In particular, the stratum

$$U_{FF} = \left\{ j : j \in U, x_{1j} = 1, x_{2j} = 1 \right\}$$

is constituted by the N_{FF} pixels classified as forest at both survey occasions, the stratum

$$U_{FN} = \left\{ j : j \in U, x_{1j} = 1, x_{2j} = 0 \right\}$$

is constituted by the N_{FN} pixels classified as forest in the first survey occasion and as non forest in the second, the stratum

$$U_{NF} = \left\{ j : j \in U, x_{1j} = 0, x_{2j} = 1 \right\}$$

is constituted by the N_{NF} pixels classified as non forest in the first survey occasion and as forest in the second, and the stratum

$$U_{NN} = \left\{ j : j \in U, x_{1j} = 0, x_{2j} = 0 \right\}$$

is constituted by the N_{NN} pixels classified as non forest at both survey occasions.

In accordance with the familiar protocol of the stratified sampling, the final sample S of fixed size n is constituted by four samples $S_{FF}, S_{FN}, S_{NF}, S_{NN}$ of size $n_{FF}, n_{FN}, n_{NF}, n_{NN}$ selected from the corresponding strata by means of SRSWOR. Accordingly, the HT estimator of the forest change is given by

$$\hat{\hat{D}}_Y = \sum_l w_l \bar{d}_{Yl} \qquad (1)$$

where the index $l = FF, FN, NF, NN$ labels the four strata, $w_l = N_l/N$ is the weight of the stratum l and

$$\bar{d}_{Yl} = \frac{1}{n_l} \sum_{j \in S_l} d_{Yj}$$

is the mean of the d_{Yj} s in the sample of pixels selected from the stratum l.

From the well-known theory of HT estimation under stratified sampling, $\hat{\bar{D}}_Y$ is unbiased with variance

$$V(\hat{\bar{D}}_Y) = \sum_l w_l^2 \frac{N_l - n_l}{N_l} \frac{S_{Yl}^2}{n_l} \tag{2}$$

where

$$S_{Yl}^2 = \frac{1}{N_l - 1} \sum_{j \in U_l} (d_{Yj} - \bar{D}_{Yl})^2$$

is the variance of the d_{Yj} s within the stratum l and

$$\bar{D}_{Yl} = \frac{1}{N_l} \sum_{j \in U_l} d_{Yj}$$

is their mean.

From Eq. (2), the precision of $\bar{D}_Y$ depends on the variability within the four strata as well as on the apportionment of the sample S to strata. As to the first source of uncertainty, it depends on how many the map changes d_{Xj} s that determine the strata match the interpreted changes d_{Yj} s. If there was a perfect correspondence between them, the within-stratum variances were 0 and (1) would estimate the forest change without error. Regarding the second source of uncertainty, a suitable option is to primarily establish a fraction f of sample observations to be assigned to the strata where changes are occurred in map data, i.e. U_{FN} and U_{NF}, in such a way that $n_c = nf$ are assigned to these strata and $n - n_c$ to the strata with no changes, i.e U_{FF} and U_{NN}. Following Olofsson et al. (2014), a large fraction f should be devoted to strata were changes are occurred.

Then, once n_c is established, a suitable option is to assign the sample units to single strata proportionally to their relative size, i.e.

$$n_{FN} = n_c \frac{N_{FN}}{N_{FN} + N_{NF}} \tag{3a}$$

$$n_{NF} = n_c \frac{N_{NF}}{N_{FN} + N_{NF}} \tag{3b}$$

$$n_{FF} = (n - n_c) \frac{N_{FF}}{N_{FF} + N_{NN}} \tag{3c}$$

$$n_{NN} = (n - n_c) \frac{N_{NN}}{N_{FF} + N_{NN}} \tag{3d}$$

4 Change estimation from spatial balance

When sampling units from a grid, there is a wide variety of spatial sampling schemes which are available besides the trivial use of SRSWOR. The achievement of SBSs, in which units are well spread throughout the survey area, has been the main target by long time. SBSs can be achieved using spatial versions of the traditional sampling schemes such as stratified or systematic sampling (e.g. Thompson 2002, Chapter 11,12) or by schemes explicitly constructed to avoid or reduce the selection of contiguous units such as the generalized random-tessellation stratified sampling by Stevens and Olsen (2004), the drawn-by-drawn sampling excluding the selection of contiguous units by Fattorini (2006), the local pivotal method of first type by Grafström et al. (2012), the spatially correlated Poisson sampling by Grafström (2012) and the doubly balanced spatial sampling by Grafström and Tillé (2013).

In this setting, the choice of effective strategies to perform forest cover estimation is a challenging issue. In this paper we based on some results from a comparison study recently performed by Fattorini et al. (2015). In presence of effective auxiliary information, as usually occurs in forest cover estimation, the D estimator is likely to outperform the HT estimator. Moreover all these schemes are likely to provide good and similar performance of the D estimator. Accordingly, the use of the OPSS seems suitable. OPSS has performance similar to the more complex explicitly-constructed spatial schemes but, contrary to these schemes, it straightforwardly provides SBS samples and can be well understood and readily planned even by non statisticians. On the other hand, when using the above-mentioned spatial schemes, the sample selection is computationally intensive and become practically impossible to apply for large populations of pixels, as those occurring in forest cover estimation. In this case, suboptimal implementations of the schemes are necessary (Grafström et al. 2014 and references therein).

Under OPSS, the population U of N pixels is partitioned into n blocks of contiguous pixels $\mathsf{U}_1, \ldots, \mathsf{U}_n$, each constituted by N/n pixels, and one pixel is randomly selected from each block. The scheme probably constitutes the first and most simple way to weaken the selection of contiguous units (Thompson 2002, Chapter 11, 12) and has a long standing in statistical literature (Breidt 1995). Under OPSS the first order inclusion probabilities of pixels are n/N, as in the case of SRSWOR, while the second-order inclusion probabilities are 0 for two pixels belonging to the same block and are n^2/N^2 otherwise.

Because under OPSS the spatial stratification is automatically performed by the scheme, ensuring a sample of pixels well spread throughout the study region, the auxiliary information is exploited at estimation level, adopting the difference (D) estimator (Särndal et al. 1992, Chapter 6). Denoting by $j_1, \ldots, j_n$ the labels of the pixels selected from $\mathsf{U}_1, \ldots, \mathsf{U}_n$, respectively, the D estimator of forest change turns out to be

$$\tilde{\bar{D}}_Y = \bar{D}_X + \frac{1}{n} \sum_{l=1}^{n} e_{j_l} \tag{4}$$

where $e_j = d_{Yj} - d_{Xj}$ is the error performed in predicting the change d_{Yj} occurred at pixel j in the reference data by means of the change d_{Xj} observed in the map data.

The D estimator is unbiased with variance

$$V(\tilde{\bar{D}}_Y) = \frac{N-n}{N} \frac{1}{n^2} \sum_{l=1}^{n} S_{El}^2 \tag{5}$$

where

$$S_{El}^2 = \frac{n}{N-n} \sum_{j \in \mathsf{U}_l} (e_j - \bar{E}_l)^2$$

is the variance of the e_j s within the stratum l and

$$\bar{E}_l = \frac{n}{N} \sum_{j \in \mathsf{U}_l} e_j$$

is their mean.

5 Analytical checking from an artificial population

The performance of the sampling strategy in which map data are exploited to partition the pixel population into change and no-change strata and the HT criterion of estimation is subsequently adopted to be compared with the sampling strategy in which spatial balance is achieved by OPSS and then the map data are exploited at estimation level by means of the D criterion. The comparison was performed on an artificial population of reference data generated from real map data.

5.1 Population

For generating a reference data population, we started from a real area, located in the Province of Nuoro within Sardinia island (Italy), constituted by a quadrat of size 400 km^2. The area was partitioned into a grid U of $N = 40,000$ pixels of size 1 ha. For this area, the Landsat classification (forest/non forest) was available for each pixel at the years 2000 and 2012, say

$$x_{00j} \text{ and } x_{12j}, \ j \in \mathsf{U}$$

from which $d_{Xj} = x_{12j} - x_{00j}$ was also available for each $j \in \mathsf{U}$. From these data, the fraction for the forest area at the years 2000 and 2012 turned out to be $\bar{X}_{00} = 0.5846$ and $\bar{X}_{12} = 0.3776$ with $\bar{D}_X = -0.2070$.

The map data at the years 2000 and 2012 were subsequently used to generate the reference data, say

$$y_{00j} \text{ and } y_{12j}, \ j \in \mathsf{U}$$

Fig. 1 **a** Changes occurred
between the years 2000–2012 on
the satellite map constituted by
40,000 quadrat pixels of size
1 ha for a quadrat area of size
20 km located in Sardinia. **b**
Changes in the reference map
artificially constructed from the
satellite map
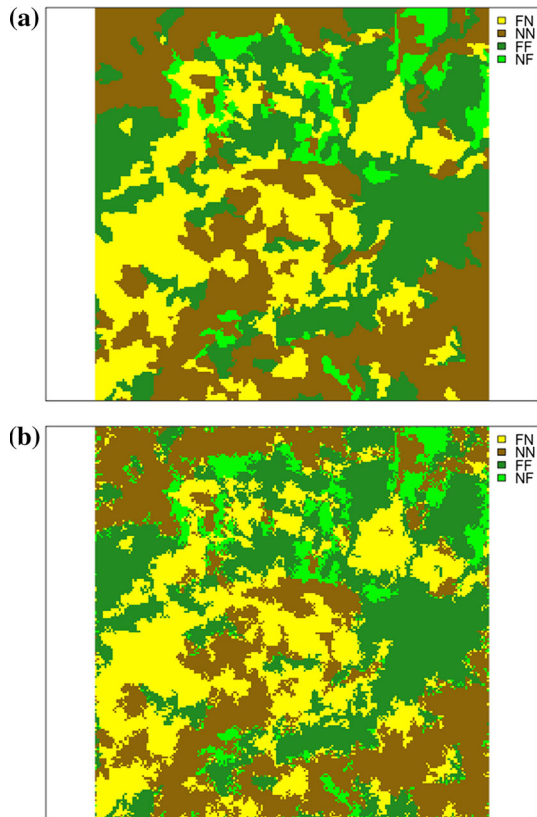
from which $d_{Yj} = y_{12j} - y_{00j}$ was also generated for each $j \in U$. To generate the reference data it was supposed that, at both years, α was the probability that a pixel classified as forest from satellite information was interpreted as forest and β the probability that a pixel classified as non-forest from satellite information was interpreted as non-forest. Thus, at both years, the interpreted forest cover at pixel j was generated from a Bernoulli random variable with parameter α if the pixel was classified as forest from satellite information and from a Bernoulli random variable with parameter $1 - \beta$ if the pixel was classified as non-forest. From some empirical investigations by Sannier et al. (2014), it is apparent that map data match reference data very well. Therefore α and β are quite high in most real situations. In this framework we presumed $\alpha = 0.90$ and $\beta = 0.85$. Finally, to avoid excessive, unrealistic fragmentation of the map, when a cluster of ten or fewer contiguous pixels of one class was completely surrounded by pixels of the other class, the cluster was assigned to the other class.

The resulting population of reference data had forest coverage $\bar{Y}_{00} = 0.5950$ and $\bar{Y}_{12} = 0.3980$ with forest change $\bar{D}_Y = -0.2060$. The correlation coefficient between the real map data and the artificial reference data were 0.80 at the year 2000 and 0.69 at the year 2012. Figure 1a shows the real changes occurred in map data from the two years while Fig. 1b shows the artificial changes occurred in reference data.

5.2 Sampling

Samples of sizes $n = 100$, 400 and 2000, corresponding to sampling fraction of 0.25, 1 and 5%, were presumed for both the strategies.

Regarding the first strategy, i.e stratified sampling joined with HT estimator, the fraction of sample observations to be assigned to the strata where a change has occurred varied from $f = 0.10$ to $f = 0.90$ with steps of size 0.10. High f values over 0.7 were in accordance with Olofsson et al. (2014), that suggest to increase the sample size for rare change classes. For any choice of f, the sample sizes within the four strata were determined in accordance to Eq. (3a–d)

Regarding the second strategy, i.e OPSS joined with D estimator, samples of size 100 were obtained by partitioning the populations into 100 quadrat blocks of 400 pixels each, and randomly selecting one pixel per block. Similarly sample of size 400 were obtained partitioning the populations into 400 quadrat blocks of 100 pixels each and samples of size 2000 were obtained partitioning the populations into 2000 rectangular blocks of 5×4 pixels each.

5.3 Performance indicators

Owing to the simplicity of the sampling schemes adopted, there was no need for simulation to determine the performance of the two sampling strategies. Both of them gave rise to design-unbiased estimators, so that their precision can be determined from their variance expressions, rather than approximated by Monte Carlo distributions, as customary in more complex cases. More precisely, the variances of the HT estimator (1) and the D estimator (4) were determined using Eqs. (2) and (5), respectively. From these quantities, the values of the *relative standard error* (RSE) were determined as the ratio $\sqrt{V}/|\bar{D}_Y|$, i.e. the square root of the variance to the absolute value of the forest change.

6 Results

Table 1 reports the percent values of RSEs for both the strategies and for any sample size. The RSE values roughly decrease at a $\sqrt{n}$-rate, ranging from 13 to 30% for $n = 100$, from 6 to 14 % for $n = 400$, and from 3 to 6% for $n = 2000$.

Table 1 Percent values of relative standard errors achieved from the two estimation strategies (Stratification + HT with different f values, OPSS + D) adopted for estimating the forest change in an artificial population of 40,000 pixels with sample sizes $n = 100, 400, 2000$

n	Stratification + HT									OPSS + D
	f 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
100	17.05	14.07	13.30	13.32	13.91	14.97	16.78	20.06	28.18	13.46
400	8.51	7.03	6.63	6.65	6.92	7.46	8.39	10.05	13.98	6.68
2000	3.76	3.08	2.91	2.91	3.04	3.28	3.70	4.45	6.22	2.88

Regarding the comparison of the two strategies, the best performance is achieved by OPSS joined with the D estimator together with the stratified sampling joined with HT estimator when the fraction of sample observations assigned to *FN* and *NF* strata is of 30 or 40%. These three strategies achieved very similar RSEs of about 13% for $n = 100$, 7% for = 400 and 3% for $n = 2000$. It is worth noting that the stratum *FN* is constituted by 10,885 pixels and the stratum NF is constituted by 2604 pixels, so that the two strata together contain about the 34% of the population. Therefore, the optimal choices of f are those nearest to the weight of the two strata in the population, i.e the best choices are akin to a proportional allocation of sample observations to strata. Unbalanced samples devoting large fractions of sample observations (80–90%) to strata where changes in map strata are occurred show the worst performance for any sample sizes.

7 Discussion and conclusions

Land cover change over time is one of the main landscape properties (e.g. Cimini et al. 2013; Hansen et al. 2013). To assess it by sampling on remotely sensed imagery, the results of our study are in favor of spatial balance of the sampled pixels or, similarly, stratification with proportional allocation of pixels to strata.

Albeit large literature from remote sensing community proposes to unbalance the sampling effort in those portions of the survey area where changes are likely to occur, from our study it is clearly apparent, on the contrary, that spatial balance—achieved in this case by the straightforward use of OPSS—provides the best performance while stratification with unbalanced assignment is generally worse, and achieves performance comparable to that provided by spatial balance only for nearly balanced assignments. The deterioration of the precision entailed by unbalanced assignments is relevant and should strongly discourage their use. From Table 1, the RSEs achieved under the maximum disequilibrium between strata ($f = 0.90$) are about two times those achieved under spatial balance or nearly proportional assignments (13 vs 28% for $n = 100$, 7 vs 14% for $n = 400$, 3 vs 6% for $n = 2000$).

These conclusions are clear: SBS with the use of map data at estimation level by means of D estimator, and stratification based on map data with proportional allocation and the use of HT estimator are both advisable in land cover change estimation and provide similar performance.

The consideration about the latter option is quite obvious. Indeed, while spatial balance ensures that the sampled pixels are evenly spread throughout the study region, the use of stratification with proportional allocation ensures that the sampled pixels are evenly allocated to the strata determined by map data. However, because these strata identify zones of changes and no changes in map data, they substantially constitute spatial strata. Therefore, the evenly apportionment of sample pixels to these strata is likely to ultimately provide a spatial balance, just like the use of an SBS scheme. At the same time, because the map data constitute very good proxies for the reference data, stratification based on map data is likely to be as effective in producing homogeneous strata, and hence enhancing the HT estimator, as the direct use of map data to predict reference data in the D estimator.

These considerations seem to justify the similar performance of the two strategies, at the same time discouraging unbalanced stratifications. Interestingly, on this issue Grafström and Tillé (2013) also provide theoretical evidence in favour of spatial balance, proving that spatial balance is the optimal strategy that minimizes the anticipated variance of the HT estimator (and, *mutatis mutandis*, of the D estimator) under a spatially auto-correlated model.

It is worth noting that very similar conclusions were achieved from a study area of the same dimension of 400 km^2 settled in Central Italy. For this area, the map data were available and showed a small forest change of about 0.2% while the weight of the two strata where changes occurred was of 10%. From the map data, the population of the reference data was artificially generated performing the same steps of Sect. 5.1. In relative terms, the results are the same as those achieved from the Sardinian case. The smallest RSE values are achieved by OPSS with D estimators. They are very similar to those obtained with stratification and HT estimator, when the allocation of sample to strata is proportional. On the other hand highly unbalanced assignments ($f = 0.90$) provide the worst RSEs that are about three times greater. Unfortunately, in absolute term the performance of the forest change estimators, even at their best, is disastrous. The smallest RSEs are of about 1800% for $n = 100$, 900% for $n = 400$, and 400% for $n = 2000$. It is indeed a well know results of spatial sampling that RSE approaches infinity as the relative size of the area under estimation approaches zero. Thus, it should be clear that very small changes cannot be estimated reliably. That invariably holds, irrespective of any guideline suggesting the use of SBSs or, that is even worse, suggesting to unbalance the sampling effort in those portions of the survey area where changes are likely to occur.

Owing to the presence of map data as highly effective proxies for reference data, we have suggested the use of difference estimator joined with OPSS to straightforwardly achieve spatial balance and efficiency. However, in presence of less effective auxiliary information, such as raw imagery data (e.g. color values) or LiDAR data, the use of spatial schemes such as the local pivotal method of first type (Grafström et al. 2012) or the doubly balanced spatial sampling (Grafström and Tillé 2013) would provide suitable alternatives to the very simple use of OPSS. Indeed, these schemes are more complex to apply, but ensure well spread samples not only with respect to the geographical coordinate but also in the space spanned by the auxiliary variables.

References

Baldauf T, Plugge D, Rqibate A, Köhl M (2009) Case studies on measuring and assessing forest degradation—monitoring degradation in the scope of REDD. FAO Forestry Department, Forest Resources Assessment Working Paper 162. Rome, 13 pp

Breidt FJ (1995) Markov chain designs for one-per-stratum sampling. Surv Methodol 21:63–70

Cimini D, Tomao A, Mattioli W, Barbati A, Corona P (2013) Assessing impact of forest cover change dynamics on high nature value farmland under Mediterranean mountain landscape. Ann Silvic Res 37:29–37

Corona P, Fattorini L, Pagliarella MC (2015) Sampling strategies for estimating forest cover from remote sensing-based two-stage inventories. For Ecosyst 2:18

Fattorini L (2006) Applying the Horvitz–Thompson criterion in complex designs: a computer-intensive perspective for estimating inclusion probabilities. Biometrika 93:269–278

Fattorini L, Corona P, Chirici G, Pagliarella MC (2015) Design-based strategies for sampling spatial units from regular grids with applications to forest surveys, land use and land cover estimation. Environmetrics 26:216–248

Grafström A (2012) Spatial correlated Poisson sampling. J Stat Plan Inference 142:139–147

Grafström A, Tillé Y (2013) Doubly balanced spatial sampling with spreading and restitution of auxiliary totals. Environmetrics 24:120–131

Grafström A, Lundström NLP, Schelin L (2012) Spatially balanced sampling through the pivotal method. Biometrics 68:514–520

Grafström A, Saarela S, Ene LT (2014) Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. Can J For Res 44:1156–1164

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2013) High-resolution global maps of 21st-century forest cover change. Science 342:850–853

Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA (2014) Good practices for estimating area and assessing accuracy of land change. Remote Sens Environ 148:42–57

Richards T, Gallego J, Achard F (2000) Sampling for forest cover change assessment at the pan-tropical scale. Int J Remote Sens 21:1473–1490

Sannier C, Mc Roberts RE, Fichet LV, Makaga EMK (2014) Using regression estimator with Landsat data to estimate proportion forest cover and net proportion deforestation in Gabon. Remote Sens Environ 151:138–148

Särndal CE, Swensson B, Wretman J (1992) Model assisted survey sampling. Springer, New York

Stehman SV (2013) Sampling strategies for forest monitoring from global to national levels. In: Achard F, Hansen MC (eds) Global forest monitoring from Earth observation. CRC Press, Boca Raton, pp 65–92

Stevens DJ, Olsen AR (2004) Spatially balanced sampling of natural resources. J Am Stat Assoc 99:262–278

Thompson SK (2002) Sampling, 2nd edn. Wiley, New York

Maria Chiara Pagliarella graduated with a Ph.D. in Quantitative Methods from University of Cassino in 2011. She was a statistical technician at the Institute Centro de Investigación Operativa (CIO) at the University "Miguel Hernandez" of Elche from 2009 to 2011. She was a post-doctoral research fellow at the Instituto de Estadística y Cartografía de la Junta de Andalucía in 2012. From 2013 to 2016 she was a research fellow in Statistics at University of Siena and University of Molise. Now she is a research fellow in Statistics at the Public Health Section, Department of Experimental Medicine of the University of Perugia. Her fields of research are: sampling theory with applications to environmental surveys, small area estimation methods, and statistics for Public Health.

Piermaria Corona is Director of the Research Centre for Forestry and Wood of CREA—Italian Council for Agricultural Research and Economics, Arezzo (Italy), since 2013. He is also Full Professor of Forest Management and Planning at the University of Tuscia, Viterbo (Italy), currently on secondment at CREA. His main research field is forest resources monitoring and management, and geomatics for natural resources assessment. He is the author of over 400 scientific and technical publications, and ten books dealing with forest monitoring issues. Professor Corona has been Dean of the School of Forest and Environmental Sciences at the Faculty of Agriculture of the University of Tuscia, from 2002 to 2009, and Director of the Department of Forest Environment and Resources of the University of Tuscia, from 2009 to 2013.

Lorenzo Fattorini is Professor of Statistics at the Department of Economics and Statistics of the University of Siena. His field of interest is sampling theory with applications to environmental surveys and ecological diversity analysis. He is Associate Editor of *Environmetrics* and *Environmental and Ecological Statistics* and from 2006 to 2009 he has been Coordinator of the Italian Working Group on Sample Survey Methodology.