Corona *et al. Forest Ecosystems* (2015) 2:18
DOI 10.1186/s40663-015-0042-7

Forest Ecosystems
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

# Sampling strategies for estimating forest cover from remote sensing-based two-stage inventories

Piermaria Corona[1*], Lorenzo Fattorini[2] and Maria Chiara Pagliarella[3]

## Abstract

**Background:** Remote sensing-based inventories are essential in estimating forest cover in tropical and subtropical countries, where ground inventories cannot be performed periodically at a large scale owing to high costs and forest inaccessibility (e.g. REDD projects) and are mandatory for constructing historical records that can be used as forest cover baselines. Given the conditions of such inventories, the survey area is partitioned into a grid of imagery segments of pre-fixed size where the proportion of forest cover can be measured within segments using a combination of unsupervised (automated or semi-automated) classification of satellite imagery and manual (i.e. visual on-screen) enhancements. Because visual on-screen operations are time expensive procedures, manual classification can be performed only for a sample of imagery segments selected at a first stage, while forest cover within each selected segment is estimated at a second stage from a sample of pixels selected within the segment. Because forest cover data arising from unsupervised satellite imagery classification may be freely available (e.g. Landsat imagery) over the entire survey area (wall-to-wall data) and are likely to be good proxies of manually classified cover data (sample data), they can be adopted as suitable auxiliary information.

**Methods:** The question is how to choose the sample areas where manual classification is carried out. We have investigated the efficiency of one-per-stratum stratified sampling for selecting segments and pixels, where to carry out manual classification and to determine the efficiency of the difference estimator for exploiting auxiliary information at the estimation level. The performance of this strategy is compared with simple random sampling without replacement.

**Results:** Our results were obtained theoretically from three artificial populations constructed from the Landsat classification (forest/non forest) available at pixel level for a study area located in central Italy, assuming three levels of error rates of the unsupervised classification of satellite imagery. The exploitation of map data as auxiliary information in the difference estimator proves to be highly effective with respect to the Horvitz-Thompson estimator, in which no auxiliary information is exploited. The use of one-per-stratum stratified sampling provides relevant improvement with respect to the use of simple random sampling without replacement.

**Conclusions:** The use of one-per-stratum stratified sampling with many imagery segments selected at the first stage and few pixels within at the second stage - jointly with a difference estimator - proves to be a suitable strategy to estimate forest cover by remote sensing-based inventories.

**Keywords:** Spatially balanced sampling; Auxiliary information; Horvitz-Thompson estimator; Difference estimator; Variance estimator; Forest monitoring

* Correspondence: piermaria.corona@entecra.it
[1]Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria - Forestry Research Centre, Viale Santa Margherita, 80 - 52100 Arezzo, Italy
Full list of author information is available at the end of the article

## Background

Deforestation and forest degradation account for nearly 20 % of global greenhouse gas emissions, more than the entire global transportation sector and second only to the energy sector. Reducing Emissions from Deforestation and Forest Degradation (REDD) is a United Nations (UN) effort to create financial value for the carbon stored in forests, offering incentives for developing countries to reduce emissions from forested lands. Monitoring systems that allow for credible measurement, reporting and verification of REDD efforts are among the most critical elements for the successful implementation of any REDD mechanism (UN-REDD 2013).

Among other issues, required data refer to updated and reliable estimates of the extent of forest cover: periodic forest cover assessments are crucial for providing benchmarks for monitoring the performance of various policies by revealing what the area of existing woodland is and whether it is increasing or declining (Corona et al. 2011; Marchetti et al. 2014). Since ground inventories cannot often be periodically performed under many large-scale conditions in tropical and subtropical countries, owing to high costs and/or lack of forest accessibility (e.g. remote and/or mountainous areas), inventories based on remote sensing imagery become compulsory. Such inventories are also mandatory for constructing historical records that can be used as a forest cover baseline.

Since quite recently, large-scale wall-to-wall remote-sensing systems are available for this purpose (see e.g. the CLASlite tool by Asner et al. 2009). However, the issue of a rigorous statistical assessment of the accuracy of the sampling strategy adopted to produce estimates is, in essence, still undervalued. As a corollary, in the REDD context, estimation must be assessed to have pre-fixed requisites of statistical accuracy (UN-REDD 2011). On this issue, it is worth noting that an objective estimation of accuracy is possible only in a design-based approach, where no assumptions are made about the population under study, in such a way that accuracy stems from the sampling strategy actually adopted to carry out estimates. Thus, accuracy is real, not assumed or modelled as in model-based approaches, where accuracy crucially depends on the model which is presumed to generate the population under study.

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) enhancements with the highest classification accuracy taken as ground truth (e.g. Hansen et al. 2013). Because visual on-screen operations are time expensive procedures, manual classification may, as a rule, be performed only for a sample of imagery segments at a first stage selection, while forest cover within each selected segment is estimated at a second stage from a sample of pixels selected within the segment. The forest cover data arising from unsupervised classification of the satellite imagery available over the whole survey area (wall-to-wall data) are likely to be good proxies of the manually classified cover data (sample data), so that they can be adopted as suitable auxiliary information (e.g. Sannier et al. 2014).

The question is how to choose the sample areas where manual classification is to be carried out. A wide variety of strategies is available to this end. The determination of the minimum -variance strategy to estimate population totals and averages is a challenging issue when the strategies are evaluated from a design-based point of view. Indeed, under design-based inference there is a lack of optimal results, in the sense that it is not possible to determine the minimum-variance strategy, as is customary in model-based approaches (e.g. Thompson 2002, Chapter 9). Accordingly, we claim no general validity about the results achieved; these should be ascribed to the conditions under which they have been obtained or extended, at most, to similar, related cases.

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 area, has been the main target for a long time. In most situations, nearby units are more similar than units far apart, thus giving a poor contribution to sample information. In these cases, the presence of spatial autocorrelation should be handled by avoiding the selection of neighboring units. From the results of a recent wide investigation (Fattorini et al. 2015), the so-called one-per stratum stratified sampling (OPSS) seems to be a suitable spatial scheme, accomplishing simplicity and efficiency, compared to other schemes investigated.

The purpose of our investigation was to compare the statistical efficiency of OPSS for selecting satellite imagery segments at a first sampling stage and pixels within the selected segments at a second sampling stage with respect to the benchmark use of simple random sampling without replacement (SRSWOR) at both stages. The comparison was performed on the basis of the actual relative root mean squared errors obtained from three artificial situations which mimic forest cover estimations.

We used Landsat imagery as a reference; since it is, by and large, the most commonly applied method to support multi-temporal delineation between forest and non-forest cover types by both manual (e.g. Townshend et al. 1995) and unsupervised automated or semi-automated (e.g. Achard et al. 2002) classification procedures and their combinations (e.g. McRoberts et al. 2014; Sannier et al. 2014).

Our presentation is organized as follows. We provide the statement of the problem. Then, we delineate the two-stage estimation of forest cover from a general point of view. Subsequently, the estimation strategies achieved using SRSWOR and OPSS, respectively, are detailed, together with problems related to variance estimation and a study is performed to compare SRSWOR and OPSS. In order to be realistic, a real map reporting the satellite classification (forest/non forest) is considered, from which the manual classification is artificially generated using assumed classification error rates. Concluding remarks are given.

It is worth noting that, even if not pursued by us, alternative criteria could be considered for selecting a sampling strategy, besides statistical efficiency. For example, cost issues could be addressed, such as when, in a multi-scene situation, image acquisition and processing costs are reduced by using multi-stage designs selecting a sample of scenes at the first stage.

## Methods

### Statement of the problem

Consider a survey area covered by a grid of $N$ rectangles of pre-fixed size, referred to as *segments* or *primary sampling units* (PSUs) by Sannier et al. (2014). Denote by $\mathsf{U}$ the population of the $N$ segments. In turn each segment is constituted by $M$ satellite pixels referred to as *secondary sampling units* (SSUs). Denote by $\mathsf{P}_j$ the population of the $M$ pixels within the $j$-th segment. As a whole, there is a target population of $N \times M$ pixels to be sampled by a two-stage scheme.

The pixel-level information consists of two dichotomous variables $x_{i,j}$ and $y_{i,j}$. The variable $x_{i,j}$ arises from forest/non-forest satellite classification and is equal to 1 if the $i$-th pixel of the $j$-th segment is classified as forest and 0 otherwise. Similarly, the variable $y_{i,j}$ arises from forest/non forest on-screen interpretation and has the value 1 if the $i$-th pixel of the $j$-th segment is interpreted as forest and 0 otherwise. While the $x_{i,j}$ variables arise from satellite spectral classes regrouped into forest and non-forest thematic classes and readily available from satellite maps for all the pixels in the population, the $y_{i,j}$ variables arise from time consuming efforts by forest experts, based on satellite imagery in combination with available very high resolution imagery and Google Earth, as well as map archives such as Bing Maps, Google Map, national maps and local maps (e.g. Sannier et al. 2014). For these reasons, the $x_{i,j}$ s are referred to as *map data* and will be used as auxiliary information in forest cover estimation, while the $y_{i,j}$s are higher quality data (e.g. more accurate data and/or with higher spatial resolution), referred to as *reference data*. These data will be taken as ground truth, given that field data cannot be collected due 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.

For each segment $j \in \mathsf{U}$, denote $x_j$ as the fraction of pixels of the $j$-th segment classified as forest from satellite information, i.e.,

$$x_j = \frac{1}{M} \sum_{i \in \mathsf{P}_j} x_{i,j}$$

Because all the $x_{i,j}$s are known, the $x_j$s are known for each $j \in \mathsf{U}$. Accordingly, their population mean

$$\bar{X} = \frac{1}{N} \sum_{j \in \mathsf{U}} x_j$$

is also known and represents the fraction of the grid area classified as forest from the satellite map. Similarly, for each segment $j \in \mathsf{U}$, denote by $y_j$ the fraction of pixels of the $j$-th segment interpreted as forest by experts, i.e.,

$$y_j = \frac{1}{M} \sum_{i \in \mathsf{P}_j} y_{i,j}$$

As stated earlier, the $y_{i,j}$s, and subsequently the $y_j$s, cannot be completely known owing to their recording costs. Accordingly their population mean

$$\bar{Y} = \frac{1}{N} \sum_{j \in \mathsf{U}} y_j$$

is unknown. It represents the fraction of grid area interpreted as forest by experts and is taken as the ground truth. Usually $\bar{Y}$ is referred to as the *forest cover*. It constitutes the target parameter to be estimated on the basis of a sampling strategy.

On the basis of empirical investigations, Sannier et al (2014) provide evidence that satellite classifications come close to resemble expert interpretation. The close matching between map and reference data suggests the use of $x_j$s as accurate and effective proxies for the $y_j$s.

### Two-stage estimation

At the first stage, consider a sampling scheme without replacement to select a sample of $n$ segments $\mathsf{S} \subset \mathsf{U}$, where $n < N$. Denote by $\pi_j$ and $\pi_{jh}$ $(h > j \in \mathsf{U})$ the first- and second-order segment inclusion probabilities induced by the first-stage scheme (see e.g. Hedayat and Sinha 1991, section 1.3). Similarly, at the second stage, consider a sampling scheme without replacement to select a sample of $m$ pixels $\mathsf{Q}_j \subset \mathsf{P}_j$, where $m < M$ for each segment $j \in \mathsf{S}$ selected at the first stage. Denote by $\tau_{i,j}$ and $\tau_{ih,j}$ $\left(h > i \in \mathsf{P}_j\right)$ the first- and second-order pixel inclusion probabilities induced at the second-stage. All these probabilities are known and previously established before sampling.

If the second stage were not performed, i.e., if all the pixels of the segments selected in the first stage were interpreted, the $y_j$s would be recorded without error and

$$\hat{\bar{Y}}_1 = \sum_{j\in \mathsf{S}} \frac{y_j}{\pi_j} \qquad (1)$$

would constitute the (virtual) one-stage Horvitz-Thompson (HT) estimator of $\bar{Y}$. From the theory of HT estimation, (1) is design-unbiased with design-based variance

$$\mathrm{V}_{\mathsf{S}}\left(\hat{\bar{Y}}_1\right) = \frac{1}{N^2} \sum_{h>j\in \mathsf{U}} \left(\pi_j\pi_h - \pi_{jh}\right)\left(\frac{y_j}{\pi_j} - \frac{y_h}{\pi_h}\right)^2 \qquad (2)$$

which represents the variance due to the first stage, i.e. the uncertainty only due to the selection of segments.

Actually, at the second stage, a sample of pixels is selected within each selected segment. From the theory of two-stage estimation (Särndal et al. 1992, Chapter 4), the two-stage HT estimator of $\bar{Y}$ is given by

$$\hat{\bar{Y}}_2 = \frac{1}{N} \sum_{j\in \mathsf{S}} \frac{\hat{y}_j}{\pi_j} \qquad (3)$$

where

$$\hat{y}_j = \frac{1}{M} \sum_{i\in \mathsf{Q}_j} \frac{y_{i,j}}{\tau_{i,j}}\ ,\ j\in \mathsf{S} \qquad (4)$$

is the HT estimator of $y_j$ obtained from the sample of pixels $\mathsf{Q}_j$ selected at the second stage for each segment $j$ selected at the first stage.

According to Särndal et al. (1992), the two-stage HT estimator (3) is design-unbiased with design-based variance

$$V\left(\hat{\bar{Y}}_2\right) = V_S\left(\hat{\bar{Y}}_1\right) + \frac{1}{N^2} \sum_{j\in \mathsf{U}} \frac{V_Q\left(\hat{y}_j\right)}{\pi_j} \qquad (5)$$

where the second term represents the increase in variance due to the second stage, i.e., the uncertainty due to the estimation of the $y_j$s in the selected segments and

$$V_Q\left(\hat{y}_j\right) = \frac{1}{M^2} \sum_{h>i\in \mathsf{P}_j} \left(\tau_{i,j}\tau_{h,j} - \tau_{ih,j}\right)\left(\frac{y_{i,j}}{\tau_{i,j}} - \frac{y_{h,j}}{\tau_{h,j}}\right)^2 \qquad (6)$$

is the variance of the HT estimator (4) for any $j\in \mathsf{U}$. Henceforth the subscript S will denote expectation and variance with respect to the first sampling stage, the subscript Q will denote expectation and variance with respect to the second sampling stage, conditional to the sample S selected in the first stage, while expectation and variance with respect to both stages are presented without a subscript.

Because the $x_j$s are known for each segment, they may be used as auxiliary information exploited at the design or estimation level. Because nothing ensures that the $x_j$s are invariably positive, they cannot be used at the design level for constructing a so-called probability-proportional-to-size scheme (Hedayat and Sinha 1991), in which the $\pi_j$s are proportional to the $x_j$s. Alternative ways to exploit the $x_j$s at the design level are the adoption of an $x$-based stratification or balancing schemes (e.g. Deville and Tillé 2004) ensuring that the sample estimate of the auxiliary mean agrees with the known population mean $\bar{X}$.

If the $x_j$s constitute good proxies for the $y_j$s, an alternative estimation strategy is to adopt a first stage scheme where the $\pi_j$s are not affected by the $x_j$s and to use the difference (D) estimator, in order to exploit auxiliary information at the estimation level by predicting the $y_j$s by means of the $x_j$s (Särndal et al. 1992, Section 6.3). Once again, if all the pixels of the segments selected at the first stage were interpreted, the $y_j$s would be recorded without error and

$$\tilde{\bar{Y}}_1 = \bar{X} + \frac{1}{N} \sum_{j\in \mathsf{S}} \frac{e_j}{\pi_j} \qquad (7)$$

would constitute the (virtual) one-stage D estimator of $\bar{Y}$, where $e_j = y_j - x_j$ denotes the error obtained from predicting $y_j$ by means of $x_j$. From the theory of D estimation, (7) is design-unbiased with a design-based variance

$$\mathrm{V}_{\mathsf{S}}\left(\tilde{\bar{Y}}_1\right) = \frac{1}{N^2} \sum_{h>j\in \mathsf{U}} \left(\pi_j\pi_h - \pi_{jh}\right)\left(\frac{e_j}{\pi_j} - \frac{e_h}{\pi_h}\right)^2 \qquad (8)$$

which represents the variance due to the first stage, i.e., the uncertainty only due to the selection of segments. On the other hand, the two-stage D estimator turns out to be

$$\tilde{\bar{Y}}_2 = \bar{X} + \frac{1}{N} \sum_{j\in \mathsf{S}} \frac{\hat{e}_j}{\pi_j} \qquad (9)$$

where $\hat{e}_j = \hat{y}_j - x_j$ is the second-stage estimate of $e_j$. From Särndal et al. (1992), the two-stage estimator (9) is design-unbiased with a design-based variance

$$\mathrm{V}\left(\tilde{\bar{Y}}_2\right) = \mathrm{V}_{\mathsf{S}}\left(\tilde{\bar{Y}}_1\right) + \frac{1}{N^2} \sum_{j\in \mathsf{U}} \frac{\mathrm{V}_{\mathsf{Q}}\left(\hat{y}_j\right)}{\pi_j} \qquad (10)$$

where the second term is the same as that of equation (5) and represents the increase in the variance due to the second stage.

Regarding the criteria for exploiting information from remote sensing at the estimation level in forest inventories, many recent papers make use of the generalized regression (GREG) estimator (e.g. Opsomer et al. 2007; Mandallaz et al. 2013; McRoberts et al. 2014). While the D estimator is used when an auxiliary variable (proxy) strictly resembles the survey variable, the GREG estimator is used when one or more auxiliary variables are strongly correlated with the survey variable. Given a close match between map and reference data, the use of the D estimator is suitable in remote sensing-based inventories. It should be noted that the D estimator is theoretically simpler than the GREG estimator, based on the fact that it is unbiased with an exact variance expression. On the other hand, the GREG estimator is not unbiased, but unbiased up to the first order of approximation, its variance is unknown and must be approximated up to the first order approximation (Särndal et al. 1993, Result 6.6.1).

Finally, for the variance estimation of both HT and D estimators, suitable solutions depend on the characteristics of the schemes adopted at the first and second stages. On this topic, it should be noted that at both stages we are dealing with the selection of spatial units (segments or pixels) from a regular grid of units, i.e., we are moving in the framework of spatial sampling. In order to reduce the selection of neighboring units (which tend to be more similar than units far apart), it is a common feature of spatial schemes that the second-order inclusion probabilities are zero or very close to zero for units that are close in distance. Moreover, explicit expressions for the second-order inclusion probabilities might be lacking for the most common spatial schemes. In these cases it is not possible to perform standard design-unbiased variance estimation and *ad-hoc* solutions should be pursued.

### Two-stage simple random sampling

The simplest way to select spatial units from a grid is to adopt SRSWOR. Accordingly, the use of SRSWOR for selecting segments at the first stage as well as pixels within segments at the second stage is considered as a benchmark. In this case, the first-stage inclusion probabilities are $\pi_j = n/N$ and $\pi_{jh} = \{n(n-1)\}/\{N(N-1)\}$ for each $h > j \in \mathsf{U}$, and $\tau_{i,j} = m/M$ and $\tau_{ih,j} = \{m(m-1)\}/\{M(M-1)\}$ for each $h > i \in \mathsf{P}_j$. From these inclusion probabilities the two-stage HT estimator (3) reduces to (Cochran 1977, Chapter 10)

$$\hat{\bar{Y}}_2 = \frac{1}{n}\sum_{j \in \mathsf{S}} \hat{y}_j \tag{11}$$

where the second-stage HT estimator (4) reduces to

$$\hat{y}_j = \frac{1}{m}\sum_{i \in \mathsf{Q}_j} y_{i,j}$$

which represents the fraction of pixels interpreted as forest out of the *m* pixels selected within segment *j*. Moreover, equation (2) reduces to

$$\mathrm{V}_{\mathsf{S}}\left(\hat{\bar{Y}}_1\right) = \frac{N-n}{N}\frac{S_y^2}{n} \tag{12}$$

where

$$S_y^2 = \frac{1}{N-1}\sum_{j \in \mathsf{U}}\left(y_j - \bar{Y}\right)^2$$

is the population variance of the $y_j$s, while equation (6) reduces to

$$\mathrm{V}_{\mathsf{Q}}\left(\hat{y}_j\right) = \frac{M-m}{M}\frac{S_j^2}{m}$$

where $S_j^2$ is the variance of the $y_{i,j}$s within segment j. Because the $y_{i,j}$s are 0-1 variables, $S_j^2$ can be rewritten as

$$S_j^2 = \frac{M}{M-1} y_j\left(1-y_j\right)$$

from which equation (6) ultimately reduces to

$$\mathrm{V}_{\mathsf{Q}}\left(\hat{y}_j\right) = \frac{M-m}{M}\frac{S_j^2}{m} = \frac{M-m}{M-1}\frac{y_j\left(1-y_j\right)}{m} \tag{13}$$

Replacing equations (12) and (13) into equation (5), the variance of $\hat{\bar{Y}}_2$ reduces to

$$\mathrm{V}\left(\hat{\bar{Y}}_2\right) = \frac{N-n}{N}\frac{S_y^2}{n} + \frac{1}{Nn}\frac{M-m}{M-1}\frac{1}{m}\sum_{j \in \mathsf{U}} y_j\left(1-y_j\right) \tag{14}$$

(see also Cochran 1977, equation 10.8).

With SRSWOR a design-unbiased estimator of variance is given by

$$\hat{V}_2 = \frac{N-n}{N}\frac{s_{\hat{y}}^2}{n} + \frac{1}{Nn}\frac{M-m}{M}\frac{1}{m-1}\sum_{j \in \mathsf{S}} \hat{y}_j\left(1-\hat{y}_j\right) \tag{15}$$

where

$$s_{\hat{y}}^2 = \frac{1}{n-1}\sum_{j \in \mathsf{S}}\left(\hat{y}_j - \hat{\bar{Y}}_2\right)^2$$

is the sample variance of the $\hat{y}_j$s (see Cochran 1977, equation 10.15 and the subsequent proof).

If the $x_j$s are used as proxies for the $y_j$s, from equation (9) the D estimator with SRSWOR reduces to

$$\tilde{\bar{Y}}_2 = \bar{X} + \frac{1}{n}\sum_{j \in \mathsf{S}} \hat{e}_j \tag{16}$$

Since under SRSWOR equation (8) reduces to

$$\mathrm{V}_{\mathsf{S}}\left(\tilde{\bar{Y}}_1\right) = \frac{N-n}{N}\frac{S_e^2}{n}$$

where

$$S_e^2 = \frac{1}{N-1}\sum_{j\in\mathsf{U}}\left(e_j-\bar{E}\right)^2$$

is the population variance of the $e_j$s and $\bar{E} = \bar{Y}-\bar{X}$ their population mean, from equation (10) the variance of $\tilde{\bar{Y}}_2$ reduces to

$$\mathrm{V}\left(\tilde{\bar{Y}}_2\right) = \frac{N-n}{N}\frac{S_e^2}{n} + \frac{1}{Nn}\frac{M-m}{M-1}\frac{1}{m}\sum_{j\in\mathsf{U}} y_j\left(1-y_j\right) \qquad (17)$$

Under SRSWOR, an unbiased estimator of variance is given by

$$\tilde{V}_2 = \frac{N-n}{N}\frac{s_{\hat{e}}^2}{n} + \frac{1}{Nn}\frac{M-m}{M}\frac{1}{m-1}\sum_{j\in\mathsf{S}} \hat{y}_j\left(1-\hat{y}_j\right) \qquad (18)$$

where

$$s_{\hat{e}}^2 = \frac{1}{n-1}\sum_{j\in\mathsf{S}}\left(\hat{e}_j-\hat{\bar{e}}\right)^2$$

is the sample variance of the $\hat{e}_j$s and $\hat{\bar{e}} = \tilde{\bar{Y}}_2-\bar{X}$ their sample mean. The proof for the unbiasedness of (18) is similar, *mutatis mutandis*, to the proof adopted by Cochran (1977) to demonstrate the unbiasedness of (15).

### Two-stage one-per-stratum stratified sampling

When sampling units from a grid, a wide variety of spatial sampling schemes is available besides SRSWOR. To obtain a SBS sample, in which units are well spread throughout the survey area, has been the main target for a long time. SBSs can be obtained using spatial versions of traditional sampling schemes such as stratified or systematic sampling (e.g. Thompson 2002, Chapters 11, 12) or by schemes explicitly constructed to avoid or reduce the selection of contiguous units such as the generalized random-tessellation stratified sampling method by Stevens and Olsen (2004), the drawn-by-drawn sampling scheme 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. Our decision of adopting OPSS at both stages is based on a study recently carried out by Fattorini et al. (2015). In the presence of effective auxiliary information, as usually occurs in forest cover estimation, all these schemes are likely to provide good and similar performance of the D estimator. Accordingly, the use of the OPSS seems suitable. Its performance is similar to the more complex explicitly-constructed spatial schemes but, in contrast 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 these spatial schemes, the sample selection is computationally intense and becomes practically impossible to apply in 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 their references).

Under OPSS, the grid $\mathsf{U}$ of $N$ segments is partitioned into $n$ blocks of contiguous segments $\mathsf{U}_1,\ldots,\mathsf{U}_n,$ each consisting of $N/n$ segments, where one segment is randomly selected from each block. Then, at the second stage, the selected segment $j\in\mathsf{S}$ is partitioned into $m$ blocks of contiguous pixels $\mathsf{P}_{j(1)},\ldots,\mathsf{P}_{j(m)},$ each consisting of $M/m$ pixels, where again one pixel is randomly selected from each block. The scheme probably constitutes the first and simplest way to weaken the selection of contiguous polygons (Thompson 2002, Chapters 11, 12) and has a long standing in statistical literature (Breidt 1995). It should be pointed out that in forest cover estimation, Sannier et al. (2014) use OPSS at the first stage to select segments, while they use simple random sampling with replacement to select pixels at the second stage within the selected segments. We have considered the use of OPSS at both stages. In this case, the first order inclusion probabilities of segments and pixels are $n/N$ and $m/M$, respectively, as is the case in SRSWOR, while those of the second order are 0 if two segments or pixels belong to the same block and are $n^2/N^2$ or $m^2/M^2$ otherwise.

From these inclusion probabilities the two-stage HT estimator coincides with the estimator (11) obtained with SRSWOR. On the other hand, with OPSS, equation (2) turns out to be

$$\mathrm{V}_{\mathsf{S}}\left(\hat{\bar{Y}}_1\right) = \frac{N-n}{Nn^2}\sum_{l=1}^{n} S_{y(l)}^2 \qquad (19)$$

where

$$S_{y(l)}^2 = \frac{1}{N/n-1}\sum_{j\in\mathsf{U}_l}\left(y_j-\bar{Y}_l\right)^2$$

is the variance of the $y_j$s within the $l$-th block of segments $\mathsf{U}_l$ and $\bar{Y}_l$ their mean, while equation (6) turns into

$$\mathrm{V}_{\mathsf{Q}}\left(\hat{y}_j\right) = \frac{M-m}{Mm^2}\sum_{k=1}^{m} S_{j(k)}^2$$

where $S_{j(k)}^2$ is the variance of the $y_{i,j}$s within block $k$ of segment $j$. Because the $y_{i,j}$s are 0-1 variables, $S_{j(k)}^2$ can be rewritten as

$$S^2_{j(k)} = \frac{M/m}{M/m-1} y_{j(k)}\left(1-y_{j(k)}\right)$$

where

$$y_{j(k)} = \frac{1}{M/m}\sum_{i\in \mathsf{P}_{j(k)}} y_{i,j}$$

is the fraction of pixels interpreted as forest within block $k$ of segment $j$. Thus, equation (6) ultimately reduces to

$$\mathsf{V}_{\mathsf{Q}}\left(\hat{y}_j\right) = \frac{1}{m^2}\sum_{k=1}^{m} S^2_{j(k)} \tag{20}$$

Replacing equations (19) and (20) in equation (5), the variance of $\hat{\hat{Y}}_2$ reduces to

$$\begin{aligned}\mathsf{V}\left(\hat{\hat{Y}}_2\right) = &\frac{N-n}{Nn^2}\sum_{l=1}^{n} S^2_{y(l)}\\ &+\frac{1}{Nn}\frac{1}{m^2}\sum_{j\in\mathsf{U}}\sum_{k=1}^{m} y_{j(k)}\left(1-y_{j(k)}\right)\end{aligned} \tag{21}$$

If the $x_j$s are used as proxies for the $y_j$s, the two-stage D estimator coincides once again with the estimator (16) obtained with SRSWOR. On the other hand, under OPSS, equation (8) reduces to

$$\mathsf{V}_{\mathsf{S}}\left(\tilde{\tilde{Y}}_1\right) = \frac{N-n}{Nn^2}\sum_{l=1}^{n} S^2_{e(l)} \tag{22}$$

where

$$S^2_{e(l)} = \frac{1}{N/n-1}\sum_{j\in\mathsf{U}_l}\left(e_j-\bar{E}_l\right)^2$$

is the variance of the $e_j$s within the $l$-th block of segments $\mathsf{U}_l$ and $\bar{E}_l$ their mean. Thus, from equation (10) the variance of $\tilde{\tilde{Y}}_2$ reduces to

$$\begin{aligned}\mathsf{V}\left(\tilde{\tilde{Y}}_2\right) = &\frac{N-n}{Nn^2}\sum_{l=1}^{n} S^2_{e(l)}\\ &+\frac{1}{Nn}\frac{1}{m^2}\sum_{j\in\mathsf{U}}\sum_{k=1}^{m} y_{j(k)}\left(1-y_{j(k)}\right)\end{aligned} \tag{23}$$

From equations (21) and (23) it is apparent that the variances of the two-stage estimators $\hat{\hat{Y}}_2$ and $\tilde{\tilde{Y}}_2$ depend on the variances of the $y_j$s or $e_j$s within the $n$ blocks partitioning the grid of segments, as well as on the variances of the $y_{i,j}$s within the $m$ blocks partitioning each segment. Because a single segment and a single pixel is selected within their corresponding blocks, it is not possible to estimate block variances from the sample information. Thus, there is no possibility to obtain design-unbiased estimators for these variances.

Conservative estimation for (21) and (23) can be attempted by using equations (15) and (18), respectively, as if SRSWOR were adopted at both stages. Because SRSWOR tends to provide greater variances than OPSS, with OPSS equations (15) and (18) tend to overestimate the actual variances (21) and (23). However, in one stage sampling, the bias induced by presuming SRSWOR when the actual scheme is OPSS, has been theoretically investigated by Mihályffy (2001); and nothing ensures that it is invariably positive. Thus, nothing ensures that (15) and (18) are invariably conservative. From tedious but conceptually simple algebra, the expectations of (15) and (18) with OPSS turn out to be

$$\begin{aligned}E\left(\hat{V}_2\right) = &\frac{N-n}{N(n-1)}\left\{\frac{N-1}{N}S^2_y - V_S\left(\hat{\hat{Y}}_1\right)\right\}\\ &+\frac{1}{N^2}\left\{\frac{N-n}{n}-\frac{M-m}{M(m-1)}\right\}\sum_{j\in\mathsf{U}} V_Q\left(\hat{y}_j\right)\\ &+\frac{1}{N^2}\frac{M-m}{M(m-1)}\sum_{j\in\mathsf{U}} y_j\left(1-y_j\right)\end{aligned} \tag{24}$$

and

$$\begin{aligned}E\left(\tilde{V}_2\right) = &\frac{N-n}{N(n-1)}\left\{\frac{N-1}{N}S^2_e - V_S\left(\tilde{\tilde{Y}}_1\right)\right\}\\ &+\frac{1}{N^2}\left\{\frac{N-n}{n}-\frac{M-m}{M(m-1)}\right\}\sum_{j\in\mathsf{U}} V_Q\left(\hat{y}_j\right)\\ &+\frac{1}{N^2}\frac{M-m}{M(m-1)}\sum_{j\in\mathsf{U}} y_j\left(1-y_j\right)\end{aligned} \tag{25}$$

where $V_S\left(\hat{\hat{Y}}_1\right)$, $V_Q(\hat{y}_j)$ and $V_S\left(\tilde{\tilde{Y}}_1\right)$ are provided by equations (19), (20) and (22), respectively.

## Analytical tests for artificial populations

The performance of the four possible strategies obtained by combining SRSWOR and OPSS with the HT and D estimators, were tested for three artificial populations.

### *Populations*

For generating the populations of $y_j$s we started from a real survey area located in central Italy, consisting of a 20 km by 20 km square area. For this area, the Landsat classifications (forest/non forest) were available for each 1 ha pixel (Fig. 1). The area was partitioned into a grid of $N = 400$ segments of 100 ha. Each segment consisted of a square grid of $M = 100$ pixels. For each segment $j$, the fraction of forest cover $x_j$ was readily calculated as the average of the $x_{i,j}$s available from the satellite map. The fraction for the whole area turned out to be $\bar{X} = 0.57$. The $x_j$s were subsequently used to generate the forest cover values $y_j$s for the three populations. We assumed that $\alpha$ was the probability that a pixel classified as forest from satellite information was interpreted as forest and $\beta$,

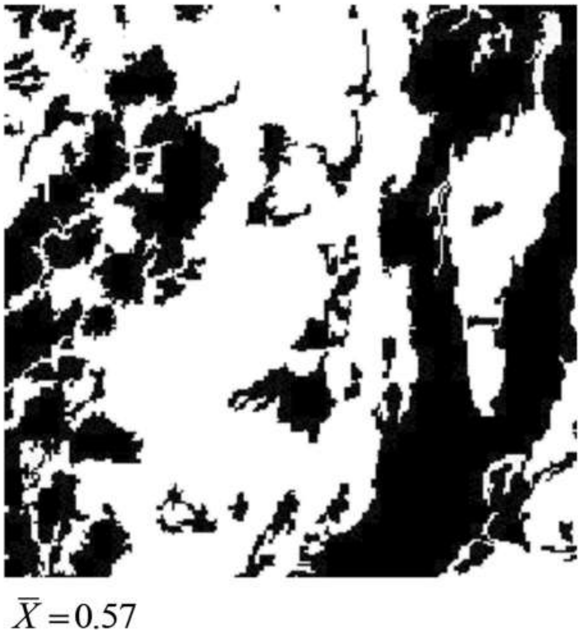


**Fig. 1** Satellite map consisting of 40,000 square pixels with a 100 m side length classified as forest (white) and non forest (black) for an area of 20 km by 20 km located in central Italy

the probability that a pixel classified as non-forest from satellite information, was interpreted as non-forest. Thus, the interpreted forest cover in the segment *j* was generated as

$$y_j = \frac{1}{M}\left\{\sum_{i\in \mathsf{P}_{j(F)}} u_{i,j} + \sum_{i\in \mathsf{P}_{j(NF)}} v_{i,j}\right\}$$

where $\mathsf{P}_{j(F)}$ and $\mathsf{P}_{j(NF)}$ were the sets of pixels of the *j*-th segment classified as forest and non-forest from satellite information and the $u_{i,j}$s and $v_{i,j}$s were independent Bernoulli random variables with parameter $\alpha$ and $1-\beta$, respectively. Clearly, for each *j*, the generated set of the *M* Bernoulli random variables gives rise to the reference values $y_{i,j}$s.

In order to obtain three different populations, the $y_j$s were generated from the $x_j$s presuming $\alpha = 0.80$ and $\beta = 0.70$ for population 1 (referred to as P1), $\alpha = 0.80$ and $\beta = 0.75$ for population 2 (P2) and $\alpha = 0.85$ and $\beta = 0.85$ for population 3 (P3). 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 populations P1, P2, and P3 (Fig. 2a, b, c) had forest cover of 0.62, 0.59, and 0.57, respectively. For each population, the graph of $y_j$s vs $x_j$s gave rise to point scatters clumped around the line $x = y$ (Fig. 3a, b, c) with $R^2$ values of 0.80, 0.92 and 0.99.

### *Sampling*

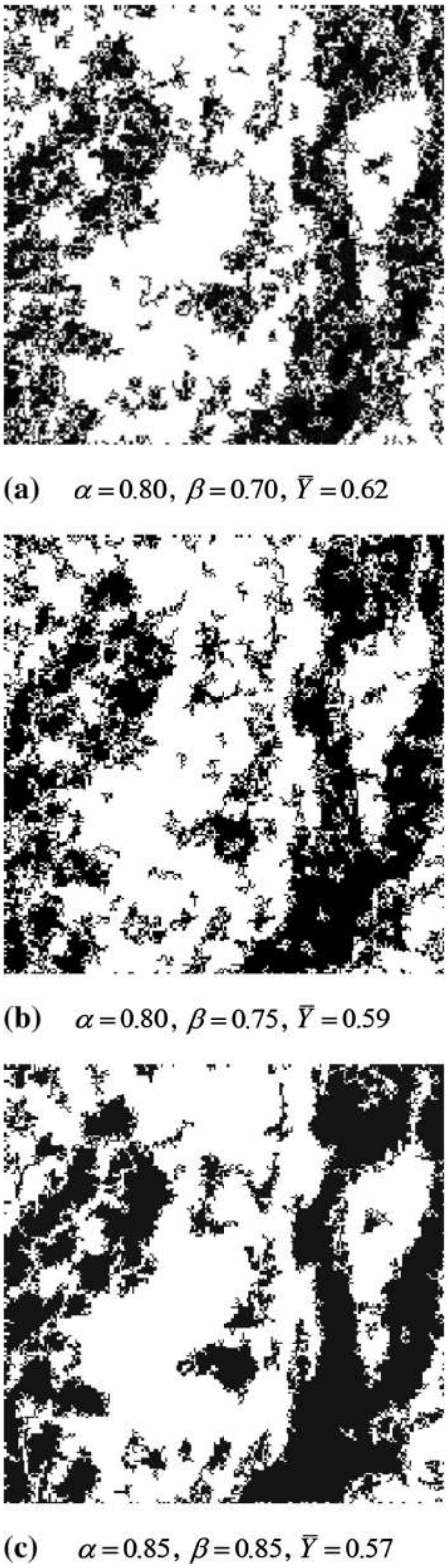
The closest matching between satellite and real classifications, also proved by Sannier et al (2014) via empirical investigations, rendered the $x_j$s accurate and effective proxies for the $y_j$s to be suitably used at the estimation level by the D estimator.

We considered SRSWOR and OPSS schemes. In first instance, these were used without exploiting auxiliary information by means of the HT estimator, followed by the exploitation of auxiliary information at the estimation level by means of the D estimator. We presumed the forest cover estimation to be as follows: for each of the four strategies, final samples of $n \times m = 100$, 400 and 2000 pixels were assumed, corresponding to 0.25 %, 1 % and 5 % sampling fractions. Samples of 100 pixels were obtained by selecting $n = 4$, 10 or 25 segments at the first stage and $m = 25$, 10 or 4 pixels at the second stage, in such a way that the total number of selected pixels turned out to be 100. Similarly, samples of 200 pixels were obtained by opting for $n = 16$, 20 or 25 and $m = 25$, 20 or 16, respectively and samples of sizes 2000 were obtained by $n = 40$, 50 or 100 and $m = 50$, 40 or 20.

### *Performance indicators*

Owing to the simplicity of the sampling schemes adopted, there was no need for simulation to determine the performance of the four sampling strategies. Each strategy gave rise to design-unbiased estimators, so that their accuracy could be determined exactly from their variance expressions, rather than approximated by Monte Carlo distributions, as is customary in more complex cases. More precisely, the variances of HT and D estimators obtained with SRSWOR were determined using equations (14) and (17), respectively, while those from OPSS were determined by equations (21) and (23). From these quantities, the values of relative standard errors (RSE) were determined as the ratio $\mathrm{SE}/\bar{Y}$ of the square root of the variance (SE) to the value under estimation, i.e., $\bar{Y}$.

For the estimation of variances, no investigations were necessary with SRSWOR, because in this case unbiased estimators were possible for the HT estimator from equation (15) and for the D estimator from equations (18). On the other hand, with OPSS the expectations of the estimators (15) and (18) were analytically determined, respectively, from equations (24) and (25). Finally the values of the relative bias (RB) were obtained as the ratio of the bias, i.e., the expectation minus the actual variance, to the actual variance. Because variances are squared quantities and as such are difficult to interpret, it is customary to estimate the relative standard error as the ratio of the square root of the estimated variance to the estimate of the population mean. As the ratio of two estimates, it is not possible to derive the expectation of the



(a) $\alpha = 0.80, \beta = 0.70, \bar{Y} = 0.62$

(b) $\alpha = 0.80, \beta = 0.75, \bar{Y} = 0.59$

(c) $\alpha = 0.85, \beta = 0.85, \bar{Y} = 0.57$

**Fig. 2** Reference maps artificially generated from the satellite map of Fig. 1, presuming three probability levels $\alpha = 0.80, 0.80, 0.85$ that a pixel classified as forest from satellite information was interpreted as forest and three probability levels $\beta = 0.70, 0.75, 0.85$ that a pixel classified as non-forest from satellite information was interpreted as non-forest

relative standard error estimator (ERSEE) analytically and therefore, we considered the first order approximation of this expected value. After some mathematical manipulation, the approximate expectation of the relative standard error estimator (AERSEE) turns out to be

$$\mathrm{AERSEE} = \mathrm{RSE}(1 + \mathrm{RB})^{1/2}$$

## Results and discussion

Table 1 presents the RSE values for each of the populations, strategies and combinations of $n$ and $m$. It also shows, in parentheses, the AERSEE values from OPSS. The results of Table 1 motivate the following inferences.

Owing to the strong correlation between map and reference data, the exploitation of map data as auxiliary information in the D estimator proves to be highly effective with respect to the HT estimator, in which no auxiliary information is exploited. With SRSWOR the decrease in RSE varies from about 20 % to 75 % and are more marked for the P3 population when the $y_j$s are maximally correlated with the $x_j$s. With OPSS, the improvement involved by using the D estimator are even more pronounced. In this case the decrease in RSE varies from about 30 % to 80 %.

The use of OPSS provides considerable improvement with respect to the use of SRSWOR. When the HT estimator is used, i.e., when no auxiliary information is exploited, the improvements are similar in each of the three populations. Decreases in RSE vary from about 10 % to 35 % and are more marked with the greatest sampling effort, i.e., of 2000 pixels. When the D estimator is used, the improvements involved when OPSS is used are even more pronounced, with decreases in RSE varying from about 20 % to 55 % and were more marked for the P3 population, when the $y_j$s were maximally correlated with the $x_j$s.

Given the repartitioning of the sampling effort between the two sampling stages, for a specific number of selected pixels $n \times m$, the performance tends to increase for the three populations and four strategies when the number of selected segments $n$ increases, with fewer pixels $m$ selected within them.

When no auxiliary information is exploited (HT), the performance tends to improve as the forest cover increases. When auxiliary information is exploited, the performance tends to improve as the correlation between auxiliary and survey variable increases.

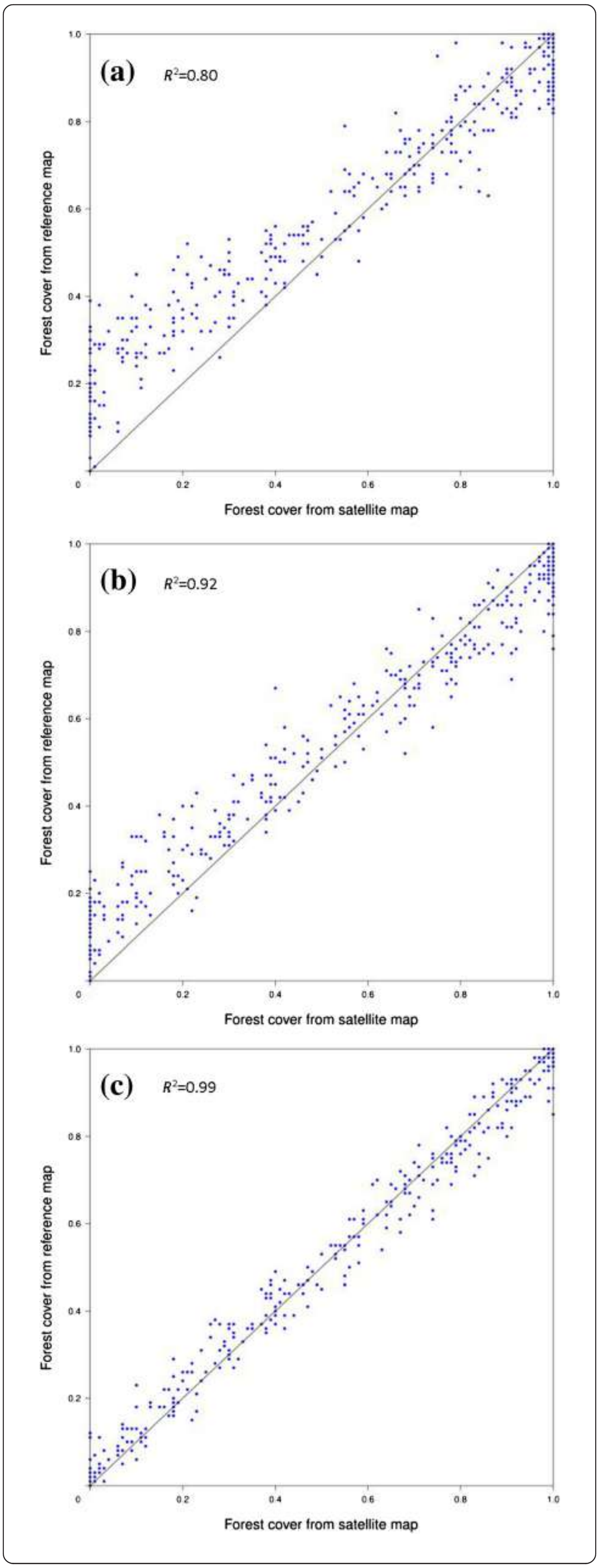


**Fig. 3** Scatter plots of forest fractions from the map of Fig. 1 vs. forest fractions from the three reference maps of Fig. 2 for the 400 segments partitioning the squared area with a side length of 20 km located in central Italy

Given the estimation of RSE with OPSS, true values are overestimated up to a maximum of 3 percentage points. In relative terms, overestimation is moderate when accuracy is poor and tends to increase as accuracy increases. Estimates of the small RSEs tend to be about twice the size of the actual values.

## Conclusions

Remote sensing from satellites is important in data collection on forest cover on a large scale and is conceived as an essential tool for REDD monitoring (UN-REDD 2013). Given this perspective, the remote sensing-based forest cover inventory procedure proposed by us proves to be suitable for any type of imagery (including that from supervised classifications or from active sensors, such as SAR interferometry or radargrammetry that can overcome eventual classification problems due to cloudiness).

From the results of our study, the use of OPSS with many image segments selected at the first stage and few pixels at the second stage - jointly with the D estimator - proves to be a suitable strategy to estimate forest cover by remote sensing-based inventories. OPSS straightforwardly provided SBSs with segments well dispersed within the survey area and pixels well spread within the selected segments. As well, the use of many segments selected at the first stage and few pixels at the second stage further increases the distribution of selected pixels onto the survey area, avoiding clumping of sample pixels within the selected segments. Finally, the D estimator exploits map data information as effective proxies of reference data. At least for the populations we investigated, the performance of this strategy was appealing. For large *n* and small *m* values, the RSEs are invariably less than 7 % even with the low sampling fraction of 0.25 % of pixels, decreasing to 1 %-1.5 % with a sampling fraction of 5 %. Because the performance of this strategy improves as the forest cover increases, better results can be expected with forest cover greater than 60 %, as frequently happens in tropical areas where coverage of about 80 % is customary (Sannier et al. 2014).

A less satisfactory issue concerns the estimation of RSE with OPSS. The proposed strategy tends to overestimate the actual values and is more marked when RSE is small. If a moderate overestimation is appealing, because it avoids the dangerous incidence of concluding that a strategy is accurate when it is not, some relevant overestimations obtained in our study are unsuitable, because they likely mask the accuracy gained by the use of OPSS jointly with the D estimator. As pointed out by Grafström (2012),

**Table 1** Values (in %) of relative standard errors obtained from the four strategies adopted in estimating remote sensing-based forest cover for the three artificial populations P1, P2 and P3 and each combination of *n* and *m*. Values in brackets represent the approximate expectations of the standard error estimators adopted with OPSS

| Population | *nxm* | *N* | *m* | SRSWOR + HT | OPSS + HT | SRSWOR + D | OPSS + D |
|---|---|---|---|---|---|---|---|
| P1 | 100 | 4 | 25 | 25.0 | 22.1 (23.2) | 14.2 | 10.3 (10.3) |
| $\bar{Y} = 0.62$ | | 10 | 10 | 18.2 | 14.4 (15.3) | 12.9 | 7.8 (8.1) |
| $R^2 = 0.80$ | | 25 | 4 | 14.2 | 9.2 (10.8) | 11.7 | 6.6 (7.3) |
| | 400 | 16 | 25 | 12.3 | 9.8 (11.5) | 7.1 | 4.7 (5.3) |
| | | 20 | 20 | 11.1 | 8.9 (10.7) | 6.5 | 4.8 (5.7) |
| | | 25 | 16 | 10.0 | 8.0 (9.9) | 5.9 | 4.7 (5.7) |
| | 2000 | 40 | 50 | 7.1 | 5.2 (7.0) | 3.4 | 2.2 (3.2) |
| | | 50 | 40 | 6.3 | 4.4 (6.6) | 3.2 | 2.4 (3.5) |
| | | 100 | 20 | 4.5 | 2.9 (5.0) | 2.8 | 1.9 (3.5) |
| P2 | 100 | 4 | 25 | 28.7 | 25.9 (27.1) | 13.4 | 8.7 (8.7) |
| $\bar{Y} = 0.59$ | | 10 | 10 | 20.4 | 16.7 (17.7) | 12.8 | 7.1 (7.4) |
| $R^2 = 0.92$ | | 25 | 4 | 15.4 | 10.2 (12.2) | 11.8 | 6.4 (7.0) |
| | 400 | 16 | 25 | 14.2 | 11.4 (13.5) | 11.8 | 4.2 (4.7) |
| | | 20 | 20 | 12.7 | 10.3 (12.4) | 6.2 | 4.5 (5.3) |
| | | 25 | 16 | 11.4 | 9.1 (11.4) | 5.7 | 4.5 (5.5) |
| | 2000 | 40 | 50 | 8.3 | 5.8 (8.2) | 3.0 | 1.9 (2.7) |
| | | 50 | 40 | 7.3 | 5.1 (7.6) | 2.8 | 2.2 (3.3) |
| | | 100 | 20 | 5.2 | 3.4 (5.7) | 2.7 | 1.8 (3.4) |
| P3 | 100 | 4 | 25 | 33.0 | 30.1 (31.6) | 11.8 | 5.2 (5.1) |
| $\bar{Y} = 0.57$ | | 10 | 10 | 22.9 | 19.1 (20.2) | 12.2 | 5.3 (5.6) |
| $R^2 = 0.99$ | | 25 | 4 | 16.7 | 11.0 (13.4) | 11.6 | 5.3 (5.8) |
| | 400 | 16 | 25 | 16.3 | 13.2 (15.6) | 5.9 | 2.6 (3.0) |
| | | 20 | 20 | 14.6 | 11.9 (14.3) | 5.6 | 3.5 (4.3) |
| | | 25 | 16 | 13.1 | 10.4 (13.0) | 5.2 | 3.8 (4.8) |
| | 2000 | 40 | 50 | 9.6 | 6.5 (9.5) | 2.3 | 1.1 (1.6) |
| | | 50 | 40 | 8.5 | 5.8 (8.6) | 2.3 | 1.6 (2.7) |
| | | 100 | 20 | 5.9 | 3.8 (6.3) | 2.5 | 1.5 (3.1) |

variance estimation is a bit tricky for spatial sampling in general and the construction of less biased variance estimators is a necessary step which calls for additional work in future investigations.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

The methodological development of the manuscript stems from the joint contribution of all the authors. Computations have been performed by MCP. All authors read and approved the final manuscript.

### Author details

[1]Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria - Forestry Research Centre, Viale Santa Margherita, 80 - 52100 Arezzo, Italy. [2]University of Siena - Department of Economic and Statistics, Piazza San Francesco, 7 - 53110 Siena, Italy. [3]University of Molise - Department of Biosciences and Territory, Contrada Fonte Lappone snc, 86090 Pesche (Isernia), Italy.

Received: 23 March 2015 Accepted: 27 May 2015
Published online: 06 June 2015

### References

Achard F, Eva HD, Stibig HJ, Mayaux P, Gallego J, Richards T, Malingreau JP (2002) Determination of deforestation rates of the world's humid tropical forests. Science 297:999–1002

Asner GP, Knapp DE, Balaji A, Paez-Acosta G (2009) Automated mapping of tropical deforestation and forest degradation: CLASlite. J Appl Remot Sens 3:1–24

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

Cochran WG (1977) Sampling Techniques. Wiley, New York

Corona P, Chirici G, McRoberts RE, Winter S, Barbati A (2011) Contribution of large-scale forest inventories to biodiversity assessment and monitoring. Forest Ecol Manag 262:2061–2069

Deville JC, Tillé Y (2004) Efficient balanced sampling: the cube method. Biometrika 91:893–912

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 Infer 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 Forest 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

Hedayat AS, Sinha BK (1991) Design and Inference in Finite Population Sampling. Wiley, New York

Mandallaz D, Brescham J, Hill A (2013) New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-based Monte Carlo approach with applications to small-area estimation. Can J Forest Res 43:1023–1031

Marchetti M, Vizzarri M, Lasserre B, Sallustio L, Tavone A (2014) Natural capital and bioeconomy: challenges and opportunities for forestry. Ann Silvicult Res 38:62–73

McRoberts RE, Liknes GC, Domke GM (2014) Using a remote sensing-based, percent tree cover map to enhance forest inventory estimation. Forest Ecol Manag 331:12–18

Mihályffy L (2001) Variance estimation for stratified samples with one unit per stratum. Hungarian Statist Rev (Special Number) 6:123–133

Opsomer JD, Breidt FG, Moisen GG, Kauermann G (2007) Model-assisted estimation of forest resources with generalized additive models. J Am Statist Assoc 102:400–416

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-Verlag, New York

Stevens DJ, Olsen AR (2004) Spatially Balanced Sampling of Natural Resources. J Am Statist Assoc 99:262–278

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

Townshend JRG, Bell V, Desch A, Havlicek C, Justice WL, Lawrence D, Skole W, Chomentowski B, Moore III, Salas W, Tucker CJ (1995) The NASA Landsat Pathfinder Humid Tropical Deforestation Project. In: Land Satellite Information in the Next Decade, ASPRS Conference Proceedings, IV-76–IV-87. ASPRS, Virginia

UN-REDD (2011) Expert meeting on assessment of forest inventory approaches for REDD+ Meeting Report. FAO, Rome

UN-REDD (2013) National Forest Monitoring Systems: Monitoring and Measurement, Reporting and Verification (M & MRV) in the context of REDD+ Activities. FAO, Rome