ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Spatial Sampling for Non-compact Patterns

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Altieri L., Cocchi D. (2021). Spatial Sampling for Non-compact Patterns. INTERNATIONAL STATISTICAL REVIEW, 89(3 (December)), 532-549 [10.1111/insr.12445].

Availability:
This version is available at: https://hdl.handle.net/11585/863155 since: 2022-02-21

Published:
DOI: http://doi.org/10.1111/insr.12445

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

14 October 2026

Spatial sampling for non-compact patterns

Linda Altieri and Daniela Cocchi
University of Bologna, Department of Statistical Sciences
via Belle Arti 41, 40126 Bologna, Italy

Abstract

The objective of spatial sampling is to collect subsets of individuals from a population in the 2-dimensional space, in order to estimate some population characteristics. Traditional sampling techniques are accordingly enriched to keep space into account. We consider sequential techniques that use weights for introducing space in the update of population units' inclusion probabilities, and propose a new weighting system that includes the spatial entropy of the study variable. Techniques only based on distances between locations perform well in the case of a compact structure. Any non-compact spatial scheme takes advantage of the involvement of spatial entropy in the sequential modification of first order inclusion probabilities.

Keywords: Environmental sampling, well spread sample, spatially correlated Poisson sampling, local pivotal method, product within distance, spatial entropy

1 Introduction

The objective of spatial sampling is to collect samples, i.e. subsets of individuals from a population, in the 2-dimensional space. Spatial sampling is strongly linked, but not limited, to environmental sampling: the data spatial location is fundamental, and sampling techniques

for environmental data are mutuated from the theory of spatial sampling. Examples can be found in biology, geography, landscape studies, forestry, and in the study of environmental dangers such as wildfires, earthquakes, polluting agents (Stevens and Olsen, 2004; Zhang and Zhang, 2012; Kermorvant et al., 2019). Other fields of application, like business surveys (Dickson et al., 2014), have been recently proposed.

In finite population inference, the design-based context aims at estimating population quantities, considered as unknown but fixed. In this case the only source of randomness comes from the selection probability of each sample, which is related to the inclusion/extraction probabilities of each population element. The equivalence of any population units for sampling translates into the exchangeability assumption with respect to the values of the variable under study, that supports simple random sampling. Information may be available for moving individual probabilities from the neutral statement of equality. Information related to space may be organized in this respect.

The link between sampling and entropy has been extensively debated in statistics (Shewry and Wynn, 1987; Lee, 2006). The search for sampling plans with high entropy is an important task in survey sampling design-based theory. In this respect, sample selection follows the idea of randomization: a sampling design should assign a non-null probability to as many samples as possible. A widely accepted measure of randomness of a sampling design is its entropy (Tillé and Haziza, 2010; Tillé and Wilhelm, 2017): a sampling design has high entropy when there is a high amount of uncertainty or surprise regarding the sample that will be selected. Poisson sampling has been identified as the maximum entropy sampling design

with fixed first order-inclusion probabilities when the sample size is random, and conditional Poisson sampling as the maximum entropy sampling design when the sample size is fixed (Hajek, 1981; Tillé, 2006; Tillé and Wilhelm, 2017). Maximum entropy sampling has been deepened in computer science, where it received important contributions (Ko et al., 1995). The theory about the entropy of sampling designs does not involve spatial sampling. Grafström (2010a) investigates the topic of entropy of sampling plans, stressing the importance of properly estimating the probabilities that enter entropy computation. His main intuition is that if good estimates of the probabilities can be found, then such entropy can be suitably approximated. An important related point is raised by Grafström (2012): high entropy is generally not a basic aim in spatial sampling, where, rather, samples that are spatially well spread in the territory are searched.

Under a different perspective, entropy is a popular heterogeneity measure for any kind of random variables. After being firstly introduced in information theory, it rapidly became popular in many applied sciences to measure the degree of heterogeneity among observations. In its original proposal, Shannon (1948) does not take space into account. A rather recent research field aims at accounting for space in entropy measures: in this spirit, a sequel of papers (Altieri et al., 2018a, 2019a,b) exploits the decomposition of bivariate distributions linked to entropy in order to quantify the contribution of spatial association to the entropy of a variable.

In what follows, we refer to the latter concept as "spatial entropy", and to the entropy of the sampling design as "sampling entropy". The two entropies need to be distinguished, as

they touch different aspects of the data. Spatial entropy refers to the spatial auto-correlation of the study variable. Sampling entropy is associated to the randomness of the samples, irrespective of the variable value.

In spatial sampling, a translation of the idea of neutrality, intended as exchangeability between units, actually implies the strong assumption of no spatial auto-correlation, which cannot be suitable in most situations. Thus, available spatial sampling methods suggest samples that are chosen based on geographical distances between data locations, irrespective of the spatial structure of the variable (Grafström, 2012; Tillé and Wilhelm, 2017). When only geographical distances are taken into account, a positive spatial auto-correlation of the values of the variable in population is implicitly assumed, which regularly decreases with distance. This is known as Tobler's Law (Tobler, 1970), that says that "near things are more related than distant things". Even if the spatially balanced sampling techniques (Benedetti et al., 2015) never explicitly consider the variable values, they produce the most efficient results when Tobler's Law holds. When data have a negative spatial auto-correlation, are spatially independent or have a weak correlation with no regular decrease, such techniques can produce inefficient samples.

Since a major challenge in spatial sampling concerns how to suitably consider data spatial auto-correlation in population, the objective of the present work is to propose a technique for spatial sampling with the ability to adapt the sample to some estimates of the spatial configuration of the study variable. A genuinely general approach to spatial sampling has no prior assumption about the spatial structure of the study variable. When the aim of

a survey is to estimate a population quantity such as the total, a good sampling plan for spatially correlated data should be able to produce similar estimates for different spatial configurations of the variable under study, i.e. the estimate and its uncertainty should not be affected by the underlying spatial structure. The spatial entropy measures proposed in Altieri et al. (2018a, 2019a,b) are employed at this regard to improve the initial sampling design by enhancing the sequential modification of the selection probabilities. This way, we not only check that spatially balanced sampling is less fruitful when Tobler's law does not hold, but also that the technique we propose produces good estimates with a mean square error that is approximately constant across spatial configurations, also for small samples. The proposal is a suggestion for real studies, where, since the spatial auto-correlation of the study variable is usually unknown, any assumptions, implicit or explicit, might lead to erroneous conclusions.

A case study is presented, which takes up the location of Swedish pine saplings (Venables and Ripley, 1997; Baddeley and Turner, 2000). The dataset, that is popular in environmental and point process studies, is a representative case of a repulsive spatial configuration due to the competition for natural resources such as soil, water, sunlight, which is frequent in applied fields such as biology, botanic, environmental studies. It is an outstanding example supporting our considerations about the available spatial sampling designs and our novel proposal, which produces efficient samples for non-compact patterns.

The paper is organized as follows. Section 2 introduces the two different ideas of entropy. Section 3 summarizes the state of the art in spatial sampling and highlights the limits of

the most recent methods. Section 3.1 presents our improvement to these proposals, which is assessed via a comparative study in Section 4. The real data application is in Section 5. Some concluding remarks are at the end of the paper.

The simulation study and all computations are implemented via the `R` software, with the help of the packages `SpatEntropy` (Altieri et al., 2018b) and `BalancedSampling` (Grafström and Lisic, 2018).

2 Spatial entropy, sampling entropy and estimation in surveys

In its simplest formulation (Cover and Thomas, 2006), the entropy of a random variable X with I categories is the expectation, under a univariate discrete probability mass function (pmf), of a random variable $I(p_X)$ known as information function:

$$H(X) = E[I(p_X)] = \sum_{i=1}^{I} p(x_i) \log\left(\frac{1}{p(x_i)}\right). \tag{1}$$

The number I of categories defines the support of (1), which may range from 0 to $\log(I)$; high values of entropy denote diversity, or surprise. A strong motivation for the popularity of $H(X)$ in applied statistics is that only $p_X = (p(x_1), \ldots, p(x_I))'$ enters the computation and not the variable values $x_1, \ldots, x_I$; therefore, entropy is computable for any type of random variable X, as long as its pmf can be computed or estimated.

2.1 Spatial entropy

Expression (1) does not explicitely consider space; several proposals have been made for spatial versions of entropy (O'Neill et al., 1988; Batty, 1974, 2010; Leibovici, 2009; Leibovici et al., 2014; Altieri et al., 2018a). The simplest effort to include space in computing entropy is based on considering a new categorical variable Z, with its pmf p_Z, whose categories are defined combining pairs of categories of X. This approach has been proposed in several biological contexts, starting from O'Neill et al. (1988). In this case, the number of different categories of Z, in other words the cardinality of the support of Z, is no longer the initial I, but a function of it, say R, which depends on whether or not order is preserved within pairs, and $H(Z)$ is computed consequently from (1). A generalization is proposed by Leibovici (2009), where "co-occurrences" are defined, consisting of couples, triples and further orders of sets of population units occurring within a predefined distance and carrying categories of the variable X. Variants of entropy measures, based on the definition of specific co-occurrences for a given distance, can be computed following this idea. The simultaneous consideration of all distances over the observation area, instead of a single predefined one, has been proposed by Altieri et al. (2018a).

2.1.1 Partial spatial mutual information

According to the setting of Altieri et al. (2018a, 2019a), Z is defined as the variable corresponding to unordered pairs of realizations of X over the observation area: $z_r = \{x_i, x'_i\}$ for $i, i' = 1, \dots I$. The variable Z has $R = \binom{I+1}{2}$ categories. A second variable W is de-

fined, classifying the Euclidean distances within the observation window according to a set of distance classes w_m, with $m = 1, \ldots, M$. A set of distance breaks $d_0, \ldots, d_M$ is fixed, where $d_0 = 0$ and d_M is the maximum possible distance inside the area; then, each class is $w_m =]d_{m-1}, d_m]$. A realization of Z, i.e. a pair, takes place at range w_m if the distance between the two units of the pair lies within the interval $]d_{m-1}, d_m]$. The variable W has pmf $p_W = (p(w_1), \ldots, p(w_M))'$, where $p(w_m)$ is the probability of any pair to fall within the mth distance range. Such setting leads to M conditional pmfs $p_{Z|w_m} = (p(z_1|w_m), \ldots, p(z_R|w_m))'$, indicating the probability of each pair to fall within distance range w_m.

The consideration of the two variables, Z and W, allows to exploit a well-known relationship of entropy theory (Cover and Thomas, 2006): the entropy of a variable may be split into the symmetric information brought by its relationship with another variable and the residual entropy due to other sources of heterogeneity

$$H(Z) = MI(Z, W) + H(Z)_W. \tag{2}$$

Since Z and W are linked to spatial information, both global residual entropy $H(Z)_W$ and mutual information $MI(Z, W)$ are spatially connotated. The first term can be renamed Spatial Mutual Information $SMI(Z, W)$, and represents the component of the entropy of Z due to its relationship with the spatial configuration. It is defined as

$$MI(Z, W) = SMI(Z, W) = \sum_{m=1}^{M} p(w_m) SPI(Z|w_m) \tag{3}$$

where each mth component $SPI(Z|w_m)$ is called Spatial Partial Information, summarizing

the behaviour of Z for each distance class w_m:

$$SPI(Z|w_m) = \sum_{r=1}^{R} p(z_r|w_m) \log \left(\frac{p(z_r|w_m)}{p(z_r)} \right). \tag{4}$$

Section 3.1 illustrates how the SPIs are useful tools in the proposal of a weighting system for spatial sampling.

2.2 Sampling entropy

The entropy of a sampling plan is a typical quantity of finite population inference computable according to (1), that possesses features that differ from the mainstream of statistical inference.

Consider a population U composed of N units, i.e. $U = \{1, \ldots, k, \ldots, N\}$. A non-random realization of X, x_i, with $i = 1, \ldots, I$ is associated to each unit. The symbol x_k identifies the value of X carried by unit k, which belongs to one of the I categories. Labelling constitutes in itself an ordering but, in the basic theory of survey sampling, labels are not related to the value of the variable under study.

Define a random variable S identifying the sample obtained without replacement; its possible realizations are $s_1, \ldots, s_J$. When the sample size $n(s)$ is random, then $J = 2^N$ (provided that $p(\emptyset) > 0$ and $p(U) > 0$); when the sample size n is fixed, then $J = \binom{N}{n} < 2^N$. Each sample s_j, with $j = 1, \ldots, J$, can be drawn with probability $p(s_j)$. A sampling design is the discrete probability distribution $p_S = (p(s_1), \ldots, p(s_j), \ldots, p(s_J))'$, with $p(s_j) \geq 0$ for all j and $\sum_{j=1}^{J} p(s_j) = 1$.

Design-based sampling theory is based on the discrete distributions induced by the se-

lection of probabilistic samples from the population U with N elements. Samples can be selected with equal or unequal probabilities, according to the design p_S. The entropy of a sampling design p_S is the entropy of the variable S, and can be written similarly to (1) as:

$$H(S) = E[I(p_S)] = \sum_{j=1}^{J} p(s_j) \log \left(\frac{1}{p(s_j)} \right), \tag{5}$$

where $0 \log 0 = 0$. Entropy $H(S)$ ranges in $[0, \log J]$, where $J \leq 2^N$ according to the sampling design. The maximum $H(S) = \log J$ is reached when all $p(s_j)$ are equal, i.e. $p(s_j) = 1/J$.

High entropy is a desirable property for a sampling design: if it occurs, a great amount of surprise is expected about the sample that can be extracted. A well-known example of a plan with low entropy is systematic sampling, whilst many popular sampling plans like simple random sampling and stratified sampling possess the desired property of maximum entropy. The spatial sampling methods used in this work enjoy the property of maximum entropy.

In sampling from finite populations, inclusion probabilities are an important feature of sampling plans and currently enter the formulae of estimators. They are population characteristics and manage the insertion of population elements in the sample. The first-order inclusion probability π_k is the probability of selecting unit k in the sample, and is the sum of the probabilities of all samples including unit k: $\pi_k = \sum_{j:k \in s_j} p(s_j)$. The second-order inclusion probability π_{kl} is the probability that two different units k and l are selected together in the sample: $\pi_{kl} = \sum_{j:\{k,l\} \subset s_j} p(s_j)$. Higher order inclusion probabilities are defined accordingly. Note that second order inclusion probabilities are the probabilities that second order co-occurrences, i.e. pairs, enter a sample.

2.3 The Horvitz-Thompson estimator

Let us suppose that the aim is to estimate the population total $t(X)$ of variable X, and, in agreement with the tradition of finite population inference, choose the Horvitz-Thompson (HT) estimator, which uses the value of the variable for the sampled units in S, x_k, and their first order inclusion probabilities π_k:

$$\widehat{t}(X) = \sum_{k \in S} \frac{x_k}{\pi_k}$$

with variance

$$V[\widehat{t}(X)] = -\frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} (\pi_{kl} - \pi_k \pi_l) \left(\frac{x_k}{\pi_k} - \frac{x_l}{\pi_l} \right)^2 .$$

Since x_k is unknown for $k \notin S$, this variance can be estimated by

$$\widehat{V}[\widehat{t}(X)] = -\frac{1}{2} \sum_{k \in S} \sum_{l \in S} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \left(\frac{x_k}{\pi_k} - \frac{x_l}{\pi_l} \right)^2 .$$

The estimator is known to be unbiased if $\pi_k > 0$ for all $k = 1, \ldots, N$, i.e. the expected value of the estimator for all possible samples under the sampling design $p(\cdot)$ equals the true total: $E_p(\widehat{t}(X)) = t(X)$. Consequently, the variance and the mean squared error

$$MSE[\widehat{t}(X)] = \frac{1}{H} \sum_{h=1}^{H} \left(\widehat{t}(X)^{(h)} - t(X) \right)^2 ,$$

with H being the number of available estimates and $\widehat{t}(X)^{(h)}$ being the hth estimated quantity, are equal.

Provided that the first- and the second-order inclusion probabilities are positive, the variance estimator gives an unbiased estimation of the mean-squared error. Usually, a normal distribution is assumed in order to quantify the uncertainty; in most sampling designs, such

assumption is asymptotically valid, and the rate of convergence depends on the sampling entropy (5). The higher the entropy, the higher the rate of convergence. Conversely, if entropy is low, i.e. the support is too small to assume that the distribution of the estimated quantity is normal. This is a reason for pursuing the maximization of the sampling entropy.

3 Entropy-based spatial sampling

In the field of balanced sampling plans with high entropy (5), most algorithms share the idea to sequentially explore the elements of the whole population, according to some order. They start by accompanying the population with a N-dimensional vector containing the first order inclusion probabilities π_k: in the first step of the algorithm, the selection probabilities are equal to the inclusion probabilities. After each decision about including a population element in the sample, the selection probabilites are sequentially modified. The vector is progressively converted into a final vector of indicators that contains only 1 or 0 values, with 1 indicating selection of the population element in the sample. The expected sample size is $E[n(s)] = n = \sum_k \pi_k$.

When space is considered, the idea of balancing may turn towards the proposal of a well spread sample. Intuitively, a sample is well spread if, for any partition of the observation area, the number of selected units per sub-area is approximately proportional to the sub-area size, i.e. if units are selected everywhere over the observation area (Stevens and Olsen, 2004; Tillé et al., 2018). We choose a number of spatially balanced sampling procedures, all

illustrated in Benedetti et al. (2017), as competitors to the present work's proposals.

The local pivotal method (LPM) (Grafström et al., 2012) is the spatial version of the pivotal method (Deville and Tillé, 1998).

The principle of LPM is to make similar units (in the sense of nearby units) compete with each other for inclusion in the sample. In the generic step of LPM, the updating of probabilities occurs between two close units k and l: unit k is chosen randomly (with equal probabilities among the units with $0 < \pi_k < 1$) and then l is chosen as its nearest neighbor (among the units with $0 < \pi_l < 1$).

Spatially Correlated Poisson sampling ($SCPS$) is the spatial version of correlated Poisson sampling, introduced by Bondesson and Thorburn (2008) and further developed by Grafström (2010b). The algorithm can be applied to any sampling design without replacement: the difference among designs lies in the choice of the initial inclusion probabilities and of the weights, that decide how sampling a specific unit is affected by the previously visited ones. If the variable under study is positively correlated in space, positive weights are attributed to units that are close in space in order to obtain the spatial spreading that is desirable for sampling. The choice of weights is computationally intriguing and not trivial: Grafström (2012) proposed a maximal weight strategy, that produces samples of fixed size and is very efficient when close units carry similar values of the study variable.

A further approach is known as Product Within Distance (PWD) (Benedetti and Piersimoni, 2017). It is based on a summary index of the matrix of distances between population units, i.e. the products of elements of the within-sample distance matrix; a tuning param-

eter can be used to increase or decrease the spread of the sample over the study region. Sampling starts with a *SRS* without replacement, then an MCMC iterative procedure repeatedly exchanges a unit included in the sample with a unit not included in the sample with a probability depending on the PWD index.

All proposals rely on the distances between population units, therefore they produce the best results when the variable is strongly spatially clustered, i.e. close units carry similar values, as stated by the popular Tobler's first law of geography.

3.1 Spatial entropy in $SCPS$

In a population of N elements, a sample of fixed size n is extracted for estimating a target quantity, say the population total $t(X)$. The same $t(X)$ and the same (non-spatial) entropies $H(X)$ and $H(Z)$ hold for any spatial configuration of the N elements of the population, while different levels of auto-correlation result in different SPI values. At this regard, in Altieri et al. (2018a, 2019a,b), four archetypical spatial configurations are identifyed, named "compact", "repulsive", "multicluster" and "random". A "good" estimator for a population synthesis like $t(X)$, "good" meaning "with a small *MSE*", ought to be found for any spatial configuration of the population.

We propose to enrich Spatially Correlated Poisson Sampling with weights that derive from the Spatial Partial Information terms (4): such weighting system takes into account the strength of auto-correlation. This new sampling procedure is fairly general, built in a very flexible way and its efficiency is not influenced by the validity of Tobler's Law. The

spatial auto-correlation of the variable is taken into account via the partial terms $SPI(Z|w_m)$ at different distance ranges, following (3) and (4). When the SPI terms are not available as population quantities, they can be estimated as proposed in the next Section.

In order to build SPI-based $SCPS$, population units are visited by the sampler according to some labelling in space. For instance, if spatial units are arranged over a grid, unit 1 can be the top-left unit, unit 2 can be at its right, or below, and so on. The method holds for any starting point and labelling criterion, as long as the distance between all pairs of units is well defined. An indicator function I_k for each visited population unit takes value 1 if the unit is sampled. When a decision is made about unit k, the remaining selection probabilities, that are initially equal to the inclusion probabilities, are updated accordingly, following a specific rule. The updating rule for $SCPS$ is based on the elements of an upper-triangular $N \times N$ matrix of weights $b_k^{(l)}$, which relate each unit k to all remaining units $l = k+1, \ldots, N$. For $k = 1$, a Bernoulli draw is performed with probability π_1; after the draw, the indicator function for that unit is $I_1 = 1$ if it is sampled, and 0 otherwise. Then, for $k = 2, \ldots, N$, the values for $I_1, \ldots, I_{k-1}$ are known and unit k is sampled with probability $\pi_k^{(k-1)}$, i.e. with a probability that was updated at the previous step when unit $k-1$ was examined. The remaining probabilities $l = k+1, \ldots, N$ are updated as:

$$\pi_l^{(k)} = \pi_l^{(k-1)} - (I_k - \pi_k^{(k-1)}) b_k^{(l)}, \tag{6}$$

and, at each step k, the probabilities of the visited units $1, \ldots, k$ leave the room to the corresponding indicator functions, until, at step N, the vector becomes $\pi_1^{(N)}, \ldots, \pi_N^{(N)} = I_1, \ldots, I_N$. The term "correlated" stresses the fact that the weights $b_k^{(l)}$ witness a form

of dependence among the inclusion probabilities. Weights should be chosen so that all probabilities, when updated, are always between 0 and 1: the criterion to meet is

$$-\min\left(\frac{1-\pi_l^{(k-1)}}{1-\pi_k^{(k-1)}}, \frac{\pi_l^{(k-1)}}{\pi_k^{(k-1)}}\right) \leq b_k^{(l)} \leq \min\left(\frac{\pi_l^{(k-1)}}{1-\pi_k^{(k-1)}}, \frac{1-\pi_l^{(k-1)}}{\pi_k^{(k-1)}}\right). \tag{7}$$

Negative weights favour the sampling of close units, since they increase the selection probability of population units that are visited after a sampled one. On the contrary, positive weights decrease the probability of population units that follow a sampled one, therefore fostering spread samples.

Each $SPI(Z|w_m)$ value is always positive, and tunes sampling neighbouring units with a strength that depends on the spatial auto-correlation of the study variable at the chosen distances. If units k and l are in the mth distance range, then the weights $b_k^{(l)}$ in (6) are

$$b_k^{(l)} = \frac{SPI(Z|w_m)}{C} \qquad \text{for } d(k,l) \in w_m \tag{8}$$

where $d(k,l)$ is the Euclidean distance between unit k and unit l; C is a normalizing constant so that each weight meets criterion (7) and $\sum_{l=k+1}^{N} b_k^{(l)} = 1$ for all k, i.e. the triangular weight matrix is row-standardized in order to obtain a fixed sample size. The easiest proposal is that the normalizing constant is the sum of the unnormalized weights: $C = \sum_{l=k+1}^{N} \widetilde{b}_k^{(l)}$ with $\widetilde{b}_k^{(l)} = SPI(Z|w_m)$ for $d(k,l) \in w_m$. Since the weights only depend on the SPI terms, they may take M different values. The special case of no auto-correlation translates into zero SPI terms and in simple random sampling of spatial units.

Working with spatial mutual information is more sophisticated than simply relying on spatial auto-correlation values. Auto-correlation is usually measured via Moran's Index

(Anselin, 1995) and is a basic way of exploring the spatial configuration of the variable. If a weighting system only based on auto-correlation values has to be constructed, it should include Moran's Index and be inversely proportional to the distance within pairs:

$$b_k^{(l)} = \frac{I_M}{C \cdot d(k,l)} \tag{9}$$

where I_M is the value of Moran's Index for the whole dataset and C is again a normalizing constant.

The present study shows that a weighting system based on Moran's Index does not perform as well as the SPI-based weights.

4 A comparative simulation study

We run a study to assess the performance of Spatially correlated Poisson sampling with SPI-based weights, compared to the version with maximal weights, to the Local Pivotal Method, to the Product Within Distance method and to simple random sampling without replacement, this seen as a benchmark. The quantity of interest is the variable total, estimated with the HT estimator. The methods' performance is evaluated via the empirical distribution of the estimates and their Mean Square Error (MSE).

A sequence of $N = 2500$ realizations is generated from a binary variable X with two alternative proportions p for outcome x_1: one is equal to 0.5, returning 1250 outcomes equal to x_1 and 1250 outcomes equal to x_0; the second one is equal to 0.25, returning 625 outcomes equal to x_1 and 1875 outcomes equal to x_0. We refer to $X_{0.5}$ and $X_{0.25}$ accordingly, the two

proportions reflecting two typical situations. Data are arranged over a 50×50 grid, each realization occurring over a square pixel of size 1. The alternative grid sizes 20×20 and 40×40 were tested, leading to analogous results. The arrangement of realizations over the grid is made following the four aforementioned spatial configurations. In the *Compact* pattern, x_1 values are assigned to the pixels located at the left part of the grid and x_0 values to pixels located at the right part. The *Random* configuration is obtained by assigning x_0 or x_1 values to pixels via simple random sampling without replacement. A *Regular* scheme is obtained by assigning x_0 values to pixels adjacent to x_1-valued pixels, and produces a perfect chessboard for $X_{0.5}$. The *Multicluster* pattern is composed by clusters, whose centroids are regularly distributed over the grid; here, the number of clusters is set equal to 16. Then, x_1 values are assigned to pixels surrounding the centroids and x_0 values to the remaining pixels. The four spatial configurations share the same $H(Z_{0.5}) = 1.04$ for the first variable generation, or $H(Z_{0.25}) = 0.86$ for the second one. All scenarios are shown in Figure 1, where x_1 values are black pixels. The four configurations differ as regards the spatial auto-correlation values, measured via Moran's Index and reported in Table 1: the *Compact* pattern follows Tobler's Law with a strong positive correlation; the *Random* configuration presents no spatial correlation; the *Regular* scheme is linked to a negative correlation; the *Multicluster* pattern shows a weak positive correlation. The evaluation conducted by Moran's Index will be improved by the use of SPI terms.

A number of sampling options is considered: MW is the acronym for $SCPS$ with maximal weight strategy, LPM stands for the Local Pivotal sampling Method, PWD for Product

Within Distance, SRS for simple random sampling without replacement, MI for $SCPS$ with correlation-based weights and SPI for $SCPS$ with SPI-based weight strategy. Several distance classifications have been tried, and $SPI4$ and $SPI70$ are here reported, where option $SPI4$ has four distance classes, while option $SPI70$ has 70 classes. This way, we propose a case with the maximum spatial detail, i.e. maximum number of classes (70 in this grid) and an alternative with a low number of classes. The choice of $SPI4$ is more computationally efficient compared to $SPI70$; moreover, it has proved to be suitable in many studies where the spatial information is present at a small scale (O'Neill et al., 1988; Altieri et al., 2018a, 2019b). The first two distance ranges are the same for the two options: they provide detailed spatial information at small distances. Class $w_1 = [0, 1]$ covers pairs of pixels whose centroids are at distance lower or equal to 1. Since in the simulated grid the pixel size is 1, class w_1 is formed by adjacent pixels, i.e. pixels sharing a border. The second class $w_2 =]1, 2]$ regards both pairs of pixels sharing a corner and pairs of pixels whose centroids are at distance 2, i.e. adjacent to a common pixel. Then, the breaks of further classes are arbitrarily chosen in both classifications. The last class is a residual one that covers all the farthest distances in the observation area: the last break is $d_{max} = 50 \times \sqrt{2} = 70.71$, i.e. the maximum distance over the square data grid. For option $SPI4$ the further classes are $w_3 =]2, 5]$ (where 5 is an exogenous choice) and $w_4 =]5, d_{max}]$. Option $SPI70$ has the most detailed distance classes for this example: all have range 1, i.e. $w_m =]m - 1, m]$ for $m = 1, \ldots, M$. The number of classes is $M = 70$, which is the integer part of d_{max}.

The chosen sampling sizes are $n_1 = N/50 = 50$, $n_2 = N/20 = 125$, and $n_3 = N/10 = 250$.

The initial inclusion probabilities are constant: $\pi_j = n_j/N$ for all units and for each sample size n_j with $j = 1, 2, 3$. A number of $H = 10000$ simulations is chosen.

Since the simulated data are arranged over the grid with one unit per cell, the selection of samples follows the grid, i.e. a sample of cells is drawn and the variable value is observed over each sampled cell. For MW, SRS and LPM we have 3×10000 samples, i.e. 10000 for each n_j: samples do not depend on p nor on the spatial configuration. Conversely, for MI, $SPI4$ and $SPI70$ the selection of the samples is different according to the proportion of x_1 and the spatial configuration of the variable. Thus, for the latter options we have $2 \times 4 \times 3 \times 10000$ samples, for two proportions, four configurations and three sample sizes. The spatial partial information values are rescaled so that they sum to 1 for each population unit, and weights are built under constraint (7).

The non-standardized $SPI(Z|w_m)$ values are shown in Figures 2 and 3. The two figures can be read according to both axes. Following the horizontal direction in Figure 2, we can see how *SPI* terms decrease with distance, with both starting value and decay depending on the strength of the spatial correlation induced by the configuration. The highest value at w_2 for the random configuration is due to the fact that, with the regular/chessboard scheme, the most homogeneous pairs, producing a low *SPI* value, are found at distance 2. Figure 3 exhibits U-shaped distributions for all distances; the shape is emphasized in the compact pattern. This is due to the fact that at larger distances a certain type of pair, i.e. $\{\text{black}, \text{white}\}$, is predominant. In both figures, the vertical axis highlights the difference across spatial configurations, and particularly how a compact pattern, where Tobler's Law

holds, is different from the other ones as regards the role of space in determining the variable outcomes. This should be accounted for in the choice of the sampled units. For more details about how to interpret *SPI* values we refer to Altieri et al. (2018a, 2019b).

4.1 Results

All sampling designs produce good results in terms of the HT estimates: the empirical distributions of 10000 values for each scenario is always concentrated around the true values $t(X_{0.5}) = 1250$ and $t(X_{0.25}) = 625$. The empirical 95% confidence intervals contain the true value in all cases. The HT estimator is unbiased, therefore a comparative performance evaluation made via the *MSE* corresponds to compare variances.

The *MSE* is shown in Table 2 for all scenarios. In the Table, a vertical line separates the results from methods available in the literature (left columns) from the ones proposed in this work (right columns). For each population proportion, sample size and spatial configuration, the best performing method is highlighted in bold.

With a focus on the existing methods (left columns), we first comment that MW, LPM and PWD have similar performances: they are mostly efficient with a compact pattern and less efficient with a multicluster configuration. For the compact pattern, MW and LPM are always very close in terms of MSE and perform better with a large sample size, while PWD is more precise with $n = 50$. In the multicluster configuration, where the spatial correlation is positive but weak, the best performing method is PWD in all scenarios. All balanced sampling methods produce the worst results in the case of negative or absent auto-

correlation: the MSEs increase abruptly wrt the case of positive auto-correlation, and are even worse than the ones deriving from SRS, not only in the random configuration but also in the repulsive pattern, where spatial auto-correlation is present but negative.

In the comparisons with the novel methods proposed in this paper (right columns), the spatially balanced methods are always the most efficient choice for the configurations with a positively autocorrelated variable; the difference reduces substantially when switching from a compact to a multicluster configuration, and in some cases (e.g. with $p = 0.5$ and $n = 125$) our methods have a lower MSE than MW and LPM. The SPI-based methods are always the most efficient in non-compact patterns: irrespective of the data proportion and sample size, they are the first and second best, and also outperform SRS in random patterns. The MI-based method, that is an alternative proposal for considering the spatial configuration of the variable while sampling, has a worse performance in all scenarios, confirming that SPI terms are more appropriate measures of the variable auto-correlation wrt Moran's Index. Moreover, our sampling approach outperforms SRS in all scenarios, which cannot be said of the balanced sampling techniques. As shown in Table 2, in many scenarios the system with 4 distance classes is enough for obtaining the best results and is also more computationally efficient than the one with 70 distance classes.

A measure of spatial balance (SBI), following Stevens and Olsen (2004), is reported in Table 3. Spatially balanced sampling techniques such as MW, LPM, PWD are built under the idea of minimizing the SBI and produce the smallest values. In all scenarios, the SPI-based methods return intermediate SBI values between spatially balanced sampling

and SRS; moreover, the SBI for SPI-based weights is smaller than the one for MI-based weights, once again supporting the use of spatial entropy as an auxiliary measure of auto-correlation in sampling. The importance of SBI is conditional on the estimator ability of producing good estimates. When examined together with the estimator efficiency, SBI supports the idea that a spatially balanced samples is the best approach when the spatial auto-correlation of the variable is strong and positive. Given the goodness of the HT estimator, well spread samples are to be preferred; in this simulation, we show that well spread samples are not always the best performing ones in terms of estimation error.

Based on the joint evaluation of the estimator efficiency and of the spatial balance index, we conclude that balanced sampling is the best option under the feeling that the variable is strongly positively correlated. In non-compact patterns, though, their MSE increases abruptly and their performance may even be worse than SRS. The magnitude of MSE values for $SPI4$ and $SPI70$ is more stable across configurations, given the data proportion p and the sample size. Thus, when the study variable is known to have a non-compact configuration, or when its spatial structure is unknown and it might be risky to make assumption, the methods proposed in this paper are a safer option, as they reduce the maximum possible estimation error and are always better than SRS.

4.2 Estimation of *SPI* terms

In simulations, the population is known and its parameters are used to assess the validity of the methods. At this regard, the procedure of SPI-based weights benefits from the

population SPI values. In real situations, though, such values are unknown and must be estimated.

In order to deepen this aspect, we built a side simulation study for SPI estimation. For each distance range w_m, the total number of pairs, say P_m, is known based on the exogenous grid size and amplitude of w_m. We proposed a stratified sampling technique using the distance ranges as strata: for each w_m, we sampled a number C_m of pairs by simple random sampling, where C_m was chosen with systematical sampling within each stratum (one out of 100), provided that each stratum contains at least 100 elements. Then, the estimates for the SPI terms were computed following (4), where probabilities are substituted by relative frequencies of the sampled pairs' categories. After marginalizing out the variable W, data can be used for estimation of probabilities for entropy $H(Z)$ (and also of $H(X)$).

Repeating this for a number of times allows to have a distribution of SPI values for each distance range. Its average values are chosen as estimators, and they are very accurate even for just 100 replicates: the difference wrt the population values is < 0.001. Such estimation procedure takes only a few minutes, therefore it does not negatively affect the computational time needed for results. In real situations, based on the available time and funding, one can choose different options for both the number of sampled pairs C_m and the number of replicates.

Computation of the HT estimator's MSEs for all scenarios leads to analogous conclusions to the ones presented in Table 2, and are not reported here.

5 The Swedish pines dataset

In this well-known example, the location of 71 Swedish pine saplings (Figure 4) is available over an area of 10 $\times$ 10 metres (Venables and Ripley, 1997; Baddeley and Turner, 2000). We discretize the area into a fine grid of 40 $\times$ 40 cells ($N = 1600$) so that each cell contains either 0 or 1 trees. The dataset looks regularly distributed as occurs in competition for resources. Moran's Index is $I_M = -0.003$, which suggests a nearly random configuration without being able to capture the repulsive behaviour of trees, well-known in the literature. Despite being the most common measure of autocorrelation, the index may induce misleading conclusions, avoidable when using entropy-based measures instead. Entropy does not reveal the type of auto-correlation, which is already known to be negative thanks to previous studies, nevertheless, when used in the sampling procedures, it is able to return samples that adapt to the data spatial structure. The global entropy values are low: $H(X) = 0.18$ and $H(Z) = 0.3$.

We choose two sample sizes: $n = 160$ and $n = 40$, and again compare several sampling designs: simple random sampling without replacement SRS, $SCPS$ with maximal weights MW, local pivotal method LPM, product within distance PWD, $SCPS$ with correlation-based weights MI, and one option for SPI, i.e. $SCPS$ with SPI-based weights. For SPI, 14 distance classes are chosen, each class with range of 1 metre expressing the maximum detail; estimation of SPI terms proceeds as in Section 4.2.

Results are reported in Table 4. Confidence intervals require estimates of the HT estimator variances. Since estimation of the variance is usually impractical in spatial sampling

(Grafström, 2012) and may lead to overestimation, we empirically estimate such variance over 1000 samples for each approach, which only takes a few minutes. As can be seen, the weighting system proposed in this work performs far better than the available alternatives in this situation, where Tobler's Law does not hold. For a sample size equal to 160 the best results are given by *SPI* and *SRS*, which outperform by far the spatial sampling methods proposed in the literature; the good performance of SRS is due to the apparent similarity of the tree pattern to a random configuration, as witnessed by Moran's Index. Our SPI-based approach, though, has a smaller standard error and is therefore more precise; moreover, results for SRS are substantially worse with a different sample size. For sample size 40, the best competitor of SPI is *LPM*, that however produces a larger confidence interval. Our SPI-based method is the only one that is reliable on a dataset with repulsive behaviour and across different sample sizes.

As for the measure of spatial balance (SBI in Table 4), the best result is achieved by PWD sampling with $n = 160$ and by LPM with $n = 40$; our SPI-based proposal gives a greater value for the index than the spatially balanced methods, but performs better than SRS and MI-based $SCPS$. The real data example supports the conclusion that a low SBI value, though desirable in general, is not necessarily to pursue, especially when the spatial configuration is non-compact. Our approach in this case leads to a less spatially balanced sample, which is nevertheless more efficient in estimating the number of trees over the area.

6 Concluding remarks

In this paper, we propose how to enhance the adaptive modification of initial first order inclusion probabilities in a spatial sampling method that enjoys the property of maximum sampling entropy. Current spatial sampling techniques (Grafström, 2010a, 2012) aim at producing spatially balanced samples, irrespective of the values of the study variable. Our conclusion is that such techniques produce the most efficient results with a positively auto-correlated study variable, but when such correlation takes a different form, or is unknown, the performance of these methods may decay fast. Our proposal, on the contrary, is able to adapt the choice of the sampled elements to the data spatial pattern, thus avoiding any risky implicit assumption.

Thanks to the simulation study, we can say that spatially balanced sampling is the most efficient in a compact pattern; by construction, it also returns a very small value for the Spatial Balance Index. The performance is still good, though decaying, for a weakly spatially correlated dataset. In non-compact patterns, such as a random or repulsive configuration, though, the efficiency of balanced sampling methods becomes much worse and they may be outperformed even by the non-spatial Simple Random Sampling. On the contrary, the SPI-based weighting system has a better performance in the latter configurations, and it also proves to perform better than an alternative weighting systems based on a standard synthesis of spatial auto-correlation such as Moran's Index. Our conclusion is that when the phenomenon under study has a non-compact configuration, or when its spatial pattern is unknown and any prior statement may lead to erroneous results, a sampling approach

which minimizes the Spatial Balance Index is not always desirable: the SPI-based sampling method may be a safer option, as it reduces the maximum possible estimation error and always performs better than non-spatial techniques such as SRS.

The real data application presented in Section 5 supports our proposed method by showing that, in a situation where a positive spatial auto-correlation cannot be assumed, an SPI-based weighting system produces a better and more precise estimate of population quantities. It also reinforces that spatial entropy may be more reliable than standard auto-correlation measures such as Moran's Index in capturing a departure from the random configuration.

The present work involves estimation of a few quantities. Estimation of the HT estimator variance is still a well-known open issue (Benedetti et al., 2015), as in spatial sampling many second order inclusion probabilities might be zero. Some proposals in the literature (Grafström, 2012) use a technique that overestimates the variance; instead, we choose to estimate it by a large number of simulations. In situations where SPI values are not available or cannot be computed, they can be estimated at a small computational cost, as proposed in Section 4.2.

The proposed methodology works for both discrete and continuous space, as shown by the application on point process data. It requires a categorical/discrete study variable, since entropy for continuous data is still unexplored, even though theoretically defined (Rényi, 1961). When continuous variables are on the fore, they may be discretized and the approach holds for any number of classes.

The application of the present work regards spatially continuous data; other spatial

datasets may be discrete, mapped over polygons or pixels. In many area-based surveys, single objects such as trees are not sampled directly: the sampling unit is a plot, which may contain several objects. As the size of the plot increases, the resulting dataset may over-represent positive spatial-correlation. For this reason, we recommend to use the finest data resolution available; in addition, since the underlying spatial structure of the variable is unknown, the use of our proposed SPI-based method may capture a potentially non-compact spatial scheme and result in better estimates of the population quantities.

Acknowledgements: this work is developed under the PRIN2015 supported project 'Environmental processes and human activities: capturing their interactions via statistical methods (EPHASTAT)' [grant number 20154X8K23] funded by MIUR (Italian Ministry of Education, University and Scientific Research).

References

Altieri, L., D. Cocchi, and G. Roli (2018a). A new approach to spatial entropy measures. *Environmental and Ecological Statistics 25*(1), 95–110.

Altieri, L., D. Cocchi, and G. Roli (2018b). *SpatEntropy: Spatial Entropy Measures*. R package version 0.1.0.

Altieri, L., D. Cocchi, and G. Roli (2019a). Measuring heterogeneity in urban expansion via spatial entropy. *Environmetrics 30*(2), e2548.

Altieri, L., D. Cocchi, and G. Roli (2019b). Advances in spatial entropy measures. *Stochastic Environmental Research and Risk Assessment 33*, 1223–1240.

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis 27*(2), 94–115.

Baddeley, A. and R. Turner (2000). Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics 42*, 283–322.

Batty, M. (1974). Spatial entropy. *Geographical Analysis 6*, 1–31.

Batty, M. (2010). Space, scale, and scaling in entropy maximizing. *Geographical Analysis 42*, 395–421.

Benedetti, R. and F. Piersimoni (2017). A spatially balanced design with probability function proportional to the within sample distance. *Biometrical Journal 59*(5), 1067–1084.

Benedetti, R., F. Piersimoni, and P. Postiglione (2015). *Sampling spatial units for agricultural surveys*. Springer.

Benedetti, R., F. Piersimoni, and P. Postiglione (2017). Spatially balanced sampling: A review and a reappraisal. *International Statistical Review 85*(3), 439–454.

Bondesson, L. and D. Thorburn (2008). A list sequential sampling method suitable for real-time sampling. *Scandinavian Journal of Statistics 35*(3), 466–483.

Cover, T. and J. Thomas (2006). *Elements of Information Theory. Second Edition*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Deville, J.-C. and Y. Tillé (1998). Unequal probability sampling without replacement through a splitting method. *Biometrika 85*(1), 89–101.

Dickson, M. M., R. Benedetti, D. Giuliani, and G. Espa (2014). The use of spatial sampling designs in business surveys. *Open Journal of Statistics 4*(5), 345–354.

Grafström, A. (2010a). Entropy of unequal probability sampling designs. *Statistical Methodology 7*(2), 84–97.

Grafström, A. (2010b). On a generalization of Poisson sampling. *Journal of Statistical Planning and Inference 140*(4), 982–991.

Grafström, A. (2012). Spatially correlated poisson sampling. *Journal of Statistical Planning and Inference 142*(1), 139–147.

Grafström, A. and J. Lisic (2018). *BalancedSampling: Balanced and Spatially Balanced Sampling*. R package version 1.5.4.

Grafström, A., N. L. Lundström, and L. Schelin (2012). Spatially balanced sampling through the pivotal method. *Biometrics 68*(2), 514–520.

Hajek, J. (1981). *Sampling from a finite population*. New York, New York: Marcel Dekker, Inc.

Kermorvant, C., F. D'Amico, and N. B. et al. (2019). Spatially balanced sampling designs for environmental surveys. *Environmental monitoring assessment 191*(8), 524–530.

Ko, C.-W., J. Lee, and M. Queyranne (1995). An exact algorithm for maximum entropy sampling. *Operations Research 43*(4), 684–691.

Lee, J. (2006). Maximum entropy sampling. *Encyclopedia of Environmetrics 4*.

Leibovici, D. (2009). *Defining spatial entropy from multivariate distributions of co-occurrences*. Berlin, Springer: In K. S. Hornsby et al. (eds.): 9th International Conference on Spatial Information Theory 2009, Lecture Notes in Computer Science 5756, 392-404.

Leibovici, D., C. Claramunt, D. L. Guyader, and D. Brosset (2014). Local and global spatio-temporal entropy indices based on distance ratios and co-occurrences distributions. *International Journal of Geographical Information Science 28*(5), 1061–1084.

O'Neill, R., J. Krummel, R. Gardner, G. Sugihara, B. Jackson, D. DeAngelis, B. Milne, M. Turner, B. Zygmunt, S. Christensen, V. Dale, and R. Graham (1988). Indices of landscape pattern. *Landscape Ecology 1*(3), 153–162.

Rényi, A. (1961). *On Measures of Entropy and Information*. University of California Press: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.

Shannon, C. (1948). A mathematical theory of communication. *Bell Dyditem Technical Journal 27*, 379–423, 623–656.

Shewry, M. C. and H. P. Wynn (1987). Maximum entropy sampling. *Journal of Applied Statistics 14*(2), 165–170.

Stevens, D. L. and A. R. Olsen (2004). Spatially balanced sampling of natural resources. *Journal of the American statistical Association 99*(465), 262–278.

Tillé, Y., M. M. Dickson, G. Espa, and D. Giuliani (2018). Measuring the spatial balance of a sample: A new measure based on the Moran's I index. *Spatial Statistics 23*, 182–192.

Tillé, Y. (2006). *Sampling algorithms*. Springer.

Tillé, Y. and D. Haziza (2010). An interesting property of the entropy of some sampling designs. *Survey Methodology 36*(2), 229–231.

Tillé, Y. and M. Wilhelm (2017). Probability sampling designs: principles for choice of design and balancing. *Statistical Science 32*(2), 176–189.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography 46*, 234–240.

Venables, W. and B. Ripley (1997). *Modern applied statistics with S-PLUS*. New York: Springer Verlag.

Zhang, J. and C. Zhang (2012). Sampling and sampling strategies for environmental analysis. *International Journal of Environmental Analytical Chemistry 92*(4), 466–478.

Table 1: Moran's Index for all scenarios.

Proportion	Spatial configuration			
	Compact	Random	Regular	Multicluster
$p = 0.5$	0.364	0	-0.011	0.038
$p = 0.25$	0.308	0	-0.009	0.040

Table 2: Estimated *MSE* of the total $t(X)$ for all scenarios and sample sizes.

$p = 0.5$							
$n = 50$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	30963	2003	1948	**1944**	16819	10035	17619
Multicluster	31792	21394	21694	**18390**	30565	25615	28050
Regular	29924	31354	32589	31112	31161	29585	**26939**
Random	30040	30786	31355	31016	25261	**24363**	24425
$n = 125$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	12026	**448**	453	467	6242	3446	11332
Multicluster	11745	5709	5646	**4734**	12026	4832	7812
Regular	11756	12165	12042	12164	14616	**7723**	8855
Random	12196	11657	12172	11780	11999	**9550**	10069
$n = 250$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	5659	145	**139**	152	2545	1135	7502
Multicluster	5609	1957	1990	**1623**	5966	2992	4867
Regular	5657	5437	5439	6270	4963	3020	**2830**
Random	5762	5592	5610	5618	4889	**4463**	4758
$p = 0.25$							
$n = 50$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	22969	1997	2022	**1985**	8966	7413	8075
Multicluster	23459	15752	15461	**12228**	22169	18119	20761
Regular	22312	24101	23621	23657	21325	**19935**	20801
Random	23028	22847	23477	21939	20001	**18883**	19939
$n = 125$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	8864	**443**	444	471	4407	2521	3416
Multicluster	8670	3734	3714	**3079**	8214	5317	6903
Regular	8693	9209	9222	9200	8700	**7234**	7829
Random	9087	9000	8800	8796	9375	8159	**8131**
$n = 250$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Compact	4288	**147**	149	159	1750	1008	2277
Multicluster	4117	1276	1304	**1083**	4071	1354	2806
Regular	4332	4829	5017	4468	4711	**4007**	4263
Random	4035	4234	4307	4080	4153	3415	**3361**

Table 3: Spatial balance index - mean and standard deviation over 10000 replicates.

$n = 50$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Mean	0.32	0.05	0.05	0.04	0.21-0.22	0.15-0.24	0.19-0.30
StDev	0.10	0.01	0.01	0.01	0.05-0.06	0.07-0.10	0.05-0.10
$n = 125$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Mean	0.31	0.05	0.05	0.03	0.22-0.23	0.13-0.19	0.18-0.28
StDev	0.06	0.01	0.01	0.01	0.04	0.03-0.07	0.04-0.08
$n = 250$							
	SRS	*MW*	*LPM*	*PWD*	*MI*	*SPI4*	*SPI70*
Mean	0.30	0.06	0.06	0.04	0.24-0.25	0.13-0.18	0.16-0.29
StDev	0.04	0.01	0.01	0.01	0.03	0.02-0.05	0.03-0.07

Approaches SRS, MW, LPM and PWD have one distribution of the SBI for each sample size over 10000 replicates. In MI- and SPI-based sampling approaches, where the sample is different based on the study variable, the SBI has a distribution for each data configuration and p: the Table reports, for each sample size, the range of results across proportions and spatial configurations.

Table 4: Swedish pines data: number of trees estimate (HT), 95% Confidence Interval (L - lower and U - upper limit) and spatial balance (SBI).

	$n = 160$				$n = 40$			
	HT	95%CI-L	95%CI-U	*SBI*	*HT*	95%CI-L	95%CI-U	*SBI*
SRS	70	63.04	76.96	0.29	40	-25.76	105.76	0.33
MW	80	62.36	97.64	0.06	120	23.96	216.04	0.06
LPM	90	50.76	129.24	0.06	80	60.36	99.64	**0.04**
PWD	90	52.76	127.24	**0.04**	40	-20.76	100.76	0.05
MI	90	52.76	127.24	0.23	200	-52.84	452.84	0.21
SPI	**70**	**68.04**	**71.96**	0.14	**80**	**62.36**	**97.64**	0.15