

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A study of variance estimation methods for systematic spatial sampling

**Geir-Harald Strand***Survey and statistics division, NIBIO, P.O.Box 115, N-1431 Ås, Norway*

H I G H L I G H T S

- Three estimators for variance in systematic spatial samples were compared.
- A correction factor based on the autocorrelation often underestimated the variance.
- A local stratified estimator and a model-based prediction both gave good estimates.

A R T I C L E I N F O

Article history:

Received 6 April 2017

Accepted 26 June 2017

Available online 8 July 2017

Keywords:

Spatial sampling

Systematic sample

Uncertainty

Variance

Land cover

Area frame survey

A B S T R A C T

An undesirable property of systematic spatial sampling is that there is no known method allowing unbiased estimation of the uncertainty of statistical estimates from these surveys. A number of alternative variance estimation methods have been tested and reported by various authors. Studies comparing these estimators are inconclusive, partly because the studies compare different sets of estimators. In this paper, three estimators recommended in recent studies are compared using a single test dataset with known properties.

The first estimator compared in this study (ST4) is based on post-stratification of the data. The second estimator (V08) is using a predetermined correction factor calculated from the spatial autocorrelation. The third estimator (MB) is a model based prediction calculated using values from the semivariogram. MB and ST4 were both found to be fairly accurate, while V08 consistently underestimated the variance in this study. V08 relies on the assumption that the autocorrelation structure in the dataset can be described using a particular exponential function. The most likely explanation of the weak result for V08 is that this assumption is violated by the empirical data used in the experiment. A better correction factor can be calculated, but the safe approach is to use MB or ST4.

E-mail address: ghs@nibio.no.<http://dx.doi.org/10.1016/j.spasta.2017.06.008>2211-6753/© 2017 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Spatial sampling is a cost-efficient way to conduct surveys for ecological and environmental studies and monitoring projects. Various sampling techniques are used when the study area is restricted in size, but national and other wide ranging surveys will often rely on systematic sampling. Many examples can be found. National forest inventories are habitually carried out using field-based systematic sampling surveys (Tomter et al., 2010; Tomppo and Tuomainen, 2010). Systematic spatial sampling provides the basis for landscape monitoring programs in Norway (Dramstad et al., 2002) and Sweden (Ståhl et al., 2011) and is used in land cover and land use surveys on a national (Strand, 2013; Aune-Lundberg and Strand, 2017) as well as a continental scale (Eurostat, 2003; Martino and Fritz, 2008). Many soil surveys also employ systematic sampling (Morvan et al., 2008).

Systematic spatial sampling is a sampling strategy with a number of favorable properties (Wang et al., 2012). It is easy to implement and there is no risk of finding sample units clustered in a few regions while other regions are left with few or no samples. In order to draw the systematic spatial sample, the population of locations must be organized as a regular frame. A starting point is drawn randomly and the rest of the sample is collected at regular intervals from this starting point. The systematic sample will result in more precise estimates than a simple random sample, in the spatial context and under commonly occurring conditions, because the sampling units are distributed more evenly across the sampled area (Bellhouse and Sutradhar, 1988; Dunn and Harrison, 1993; D'Orazio, 2003; Ambrosio et al., 2004).

The systematic sample is in particular preferable as a sampling method when nearby sampling units show a high degree of positive correlation (Cochran, 1977). This was demonstrated by Flores et al. (2003) who compared the relative efficiency of systematic sampling to simple random sampling from populations with known properties. The study demonstrated that systematic sampling was more efficient than simple random sampling and showed that the improvement in efficiency was related to sampling distance. The relative efficiency of systematic sampling was higher when the sampling distance was short and lower when the sampling distance increased. The change in relative efficiency was closely related to the spatial autocorrelation.

An undesirable property of systematic sampling is that there is no known method allowing unbiased estimation of the uncertainty in these surveys. The higher precision achieved by systematic sampling may therefore go unnoticed. The reason for this shortcoming is found in the systematic sample design, where the population – at least in theory – is divided into a number of partitions. Each partition consists of the population elements included in the sample when a particular starting point is selected. There is a finite set of starting points representing a finite set of partitions (Madow and Madow, 1944). Each and every population element is assigned to one (and only one) partition. When a partition is included in the sample, then every population element in this partition is included (Thompson, 2002 pp. 129–131). A simple example is illustrated in Fig. 1 where a population of grid cells is divided into four partitions labeled A, B, C and D.

Systematic sampling is (usually) limited to drawing a single partition by choosing a single starting point. This is equivalent to a sample size of $n = 1$ partitions. Ordinary variance estimation methods require a denominator of $n - 1$ and can therefore not be applied (Thompson, 2002).

The conservative approach for handling uncertainty in a systematic sample is to calculate the variance using the estimators intended for simple random sampling (Milne, 1959; Cochran, 1977; Wolter, 1984, 2007). This is usually a safe approach and will in certain situations be both acceptable and commendable, but has a tendency to overestimate the variance (McRoberts et al., 2016). A large number of alternative, more or less biased, variance estimation methods have been tested and reported by various authors (Matèrn, 1947, 1960; Wolter, 2007; Gallego and Delincé, 2010; Aubry and Debouzie, 2000; Dunn and Harrison, 1993; D'Orazio, 2003; Opsomer et al., 2012).

C	D	C	D	C	D
A	B	A	B	A	B
C	D	C	D	C	D
A	B	A	B	A	B
C	D	C	D	C	D
A	B	A	B	A	B

Fig. 1. A population of 36 tiles divided into partitions for a sampling interval of $d = 2$ (every second tile in both cardinal directions). The result is $d^2 = 2^2 = 4$ partitions (A, B, C and D). The systematic random sampling approach is to randomly choose one of these partitions.

Several of the estimators were compared in two papers by [Aune-Lundberg and Strand \(2014\)](#) and [McGarvey et al. \(2016\)](#). The results from these two studies are inconclusive. [Aune-Lundberg and Strand \(2014\)](#) found post-stratification using small strata to be the most efficient approach, while [McGarvey et al. \(2016\)](#) found an estimator using a correction factor based on the measurement of spatial autocorrelation to be most efficient. Unfortunately, the two studies do not include each other's recommended estimator. They were also carried out in two different environments with different data qualities: [Aune-Lundberg and Strand \(2014\)](#) used real-world data covering an entire country while [McGarvey et al. \(2016\)](#) used a large, synthetic dataset.

A recent addition to the literature on variance estimation methods for systematic spatial sampling is ([Brus and Saby, 2016](#)) who compared five estimation methods. In their study, they found that the method recommended by [McGarvey et al. \(2016\)](#) performed less well. Instead, they found that a model-based prediction using values extracted from the semivariogram was the best approach. They also found that post-stratification using a method fairly similar, but not identical, to the one proposed by [Aune-Lundberg and Strand \(2014\)](#) was a reasonable approximation, and in some situations also better than the model-based approach.

The objective of the current study is to compare the three estimators recommended respectively by [Aune-Lundberg and Strand \(2014\)](#), [McGarvey et al. \(2016\)](#) and [Brus and Saby \(2016\)](#). This is done by applying all three estimators to samples drawn from a single, large dataset with known properties.

2. Material and methods

2.1. Material

The data used in the study were compiled from a detailed national land cover map of Norway. The minimum mapping unit is around 1.5 hectare with a geometric accuracy of 20 m. The map is available on the internet (<http://kilden.nibio.no>—last accessed on June 23rd 2017). The study area was the entire Norwegian mainland, totally 324,099 km². The study used a land cover classification with seven land cover classes listed in [Table 1](#).

The land cover was partitioned into quadratic one square kilometer tiles based on the standardized statistical grid for Norway provided by Statistics Norway ([Strand and Bloch, 2009](#)), resulting in a population consisting of $N = 350,514$ regular tiles. By using a GIS overlay function, the acreage of each land cover class was calculated for each tile (grid cell) as percent (%) of the tile. Since the entire population was known, key statistics (mean and variance of the coverage per tile) could be computed for every land cover class. The result is found in [Table 1](#).

Table 1

Descriptive statistics (sum, population mean and population variance) for the seven land use/land cover types in the gridded version of the national land use/land cover map AR50. $N = 350,514$ grid cells.

Land cover class	N	Sum (km ²)	Mean μ (%)	Variance σ^2
1 Built-up land	350,514	1,859.25	0.5304	21.049
2 Agriculture	350,514	12,658.59	3.6114	137.354
3 Forest	350,514	126,113.46	35.9796	1340.331
4 Open land	350,514	140,148.26	39.9836	1714.754
5 Mire	350,514	21,722.85	6.1974	161.123
6 Snow/Ice	350,514	3,038.19	0.8668	59.342
7 Water	350,514	18,559.31	5.2949	200.692

2.2. Method

Systematic sampling entails splitting the population elements (in this case tiles) into a number of discontinuous partitions following a regular design as illustrated in Fig. 1. The number of partitions depends on the sampling interval. A sampling interval of two will result in four partitions, each containing every second tile in both cardinal directions (as in Fig. 1). A sampling interval of three results in nine partitions etc. In general, a sampling interval of d results in $k = d^2$ partitions. Systematic sampling amounts to selecting one of the partitions and the sample itself consists of all the population elements in the selected partition.

Since a complete dataset for the entire population was available, the three variance estimators as well as other relevant parameters were calculated for every partition in the population. For each sampling interval d , the population was subdivided into partitions by randomly choosing a block of d by d tiles to initiate the partition. Each of the $k = d^2$ tiles in this block was used as the starting point for one partition, by including every d^{th} grid cell in both cardinal directions from the initial tile. Each partition is a possible systematic random sample, and the k partitions defined by the exercise included all the possible partitions in the population (based on the standardized grid and a sampling intensity using d as the sampling interval).

The study first compared variance estimators for all seven land cover classes at the sampling interval 10 km. This is identical to the sampling interval used in Aune-Lundberg and Strand (2014). The same variance estimators were also compared for one particular land cover class over all sampling intervals in the range 2–20 sampling units (in this case representing distances from 2 up to 20 km). The land cover class Mire and peat bog was used as the example for the latter part of the study. The spatial distribution of the land cover class Mire and peat bog is shown as a thematic map in Fig. 2.

With seven land cover classes and d^2 partitions for each sampling interval in the range 2–20 km, the total number of partitions in the study was 20,083. The $k = d^2 = 100$ candidate partitions for each land cover class at the sampling interval $d = 10$ km consisted of approximately 3505 tiles each. For the land cover class Mire and peat bog in particular, all candidate partitions were used. This approach gave $\sum_{d=2}^{20} d^2 = 2869$ partitions of variable sizes for comparing the estimators at different sampling intervals.

The (exact) systematic sample variance ($VAR(\bar{x})_{SYS}$) for a certain land cover type at a certain sampling interval was determined empirically from the k candidate partitions at that particular sampling interval.

$$VAR(\bar{x})_{SYS} = \frac{1}{k} \sum_{j=1}^k (\hat{x}_j - \bar{x})^2 \text{ where}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\hat{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i.$$

N is the total number of tiles in the entire population, k is the number of candidate partitions and n_j is the number of tiles in partition j . The calculation is based on the assumption that \hat{x}_j is an unbiased estimator of \bar{x} . This assumption was verified empirically.

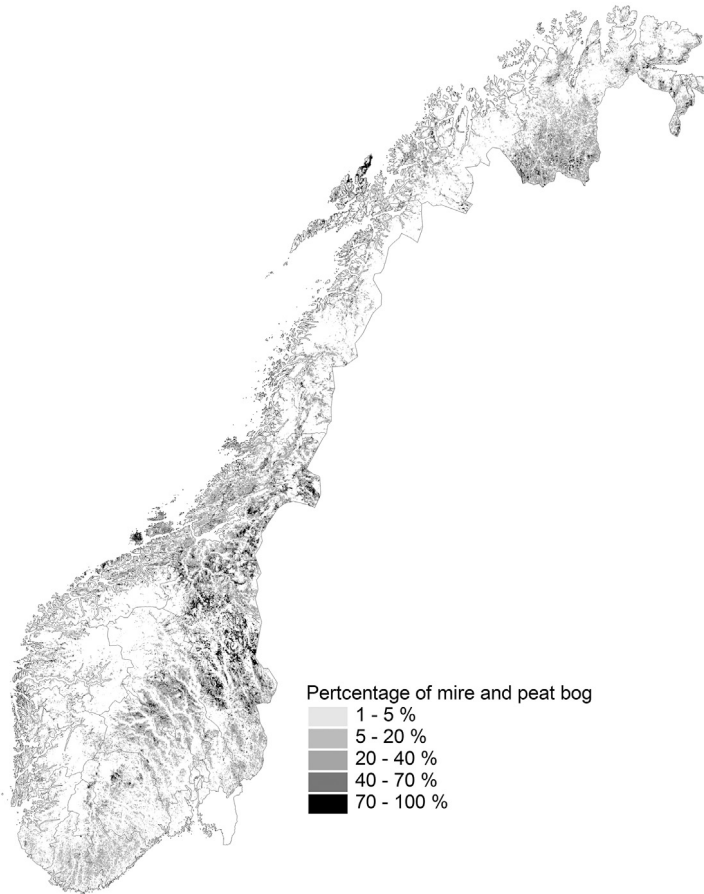


Fig. 2. Distribution of the land cover class Mire and peat bog in Norway.

Several estimators of variance were calculated for every sample. The estimators were $VAR(\bar{x})$ SRS, $VAR(\bar{x})$ ST4, $VAR(\bar{x})$ V08 and $VAR(\bar{x})$ MB. $VAR(\bar{x})$ SRS is the naïve estimator calculated by treating the sample as a simple random sample. This estimator was included as a reference. $VAR(\bar{x})$ ST4 is the estimator recommended by [Aune-Lundberg and Strand \(2014\)](#) and $VAR(\bar{x})$ V08 is the estimator found to be most efficient by [McGarvey et al. \(2016\)](#). $VAR(\bar{x})$ MB is the model-based estimator proposed by [Brus and Saby \(2016\)](#)

The naïve estimator was calculated as

$$VAR(\bar{x}) \text{ SRS} = \frac{1}{n} s_x^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \hat{x})^2$$

where n is the sample size and

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

This estimator is only correct when the observations are independent and identically distributed (IID), an assumption often violated in spatial samples due to the presence of spatial autocorrelation. The estimator was included in the study as a reference only.

$VAR(\bar{x})ST4$ is a local estimator calculated using non-overlapping strata where each stratum is a two by two units neighborhood of sample tiles. The estimation method is the approach used in stratified random sampling

$$VAR(\bar{x})ST4 = \sum_{j=1}^b w_j^2 \frac{s_j^2}{q_j} \frac{(N_j - k_j)}{N_j}$$

where b is the number of strata, q_j is the number of cases in stratum j (mostly four) and s_j^2 is the observed variance inside stratum j calculated as

$$s_j^2 = \frac{1}{q_j} \sum_{i=1}^{q_j} (x_i - \hat{x}_j)^2.$$

Notice that the divisor is set to q_j (and not to $q_j - 1$) since the objective is to calculate the actual variance in the stratum, and not to find an unbiased estimate of the population variance. Furthermore, N_j is the population size in stratum j (in our case always set to $d^2 \times q_j$ since each tile in the sample “represents” d^2 tiles) and:

$$w_j = \frac{N_j}{N} \quad \text{or (if } N \text{ is unknown) } w_j = \frac{q_j}{n}$$

where N is the total number of tiles in the population and n is the sample size.

$VAR(\bar{x})V08$ is an estimator originally studied by Cochran (1946) and retrieved by McGarvey et al. (2016) from Wolter (2007, p. 302) where it is called $v8$. The estimator is calculated as

$$VAR(\bar{x})V08 = \frac{s_x^2 (N - n)}{n (N - 1)} \left[1 + \frac{2}{\ln(\hat{\rho}_d)} + \frac{2}{(\hat{\rho}_d^{-1} - 1)} \right] \quad \text{if } \hat{\rho}_d > 0$$

$$VAR(\bar{x})V08 = \frac{s_x^2 (N - n)}{n (N - 1)} \quad \text{if } \hat{\rho}_d \leq 0$$

where s_x^2/n is the naïve estimator, identical to $VAR(\bar{x})SRS$ above.

The parameter $\hat{\rho}_d$ used in $VAR(\bar{x})V08$ is the estimated spatial autocorrelation at distance d .

Spatial autocorrelation is the phenomenon that observations taken close to each other tend to be more similar than observations taken further apart (Legendre, 1993) and that this effect is a function closely related to distance. A global (non-spatial) measurement of autocorrelation is the intra-class correlation between all pairs of observations in a sample (or a population).

$$\rho = \frac{1}{n(n-1)s^2} \sum_i^n \sum_{j \neq i}^n (x_i - \bar{x})(x_j - \bar{x}).$$

The intra-class correlation is identical to the more familiar bivariate correlation coefficient, but involves only one variable. The range is the same (from +1 to -1, with 0 representing no correlation). Spatial autocorrelation is intra-class correlation calculated as a function of separation distance. The intra-class correlation is, for this purpose, calculated for all the pairs of observations separated by the sampling distance d or with separation distance within a particular range $h = \{d_1..d_2\}$. In the last case, h is called a lag.

$$\rho_h = \frac{1}{m(m-1)s_x^2} \sum_i^m \sum_{j \neq i}^m (x_i - \bar{x})(x_j - \bar{x})$$

where m is the number of pairs $\{x_i, x_j\}$ separated by a distance within lag h . $\hat{\rho}_h$ is the estimate of ρ_h calculated using the available data from the sample.

The spatial autocorrelation function is a graph showing ρ_h along the vertical axis, as a function of the separation distance h (represented along the horizontal axis). This graph is known as a correlogram, and was used to visualize the autocorrelation function. The empirical autocorrelation $\hat{\rho}_h$ was calculated for each sample using the sample data and used in the calculation of $VAR(\bar{x})$ V08.

The variance estimator $VAR(\bar{x})$ MB is explained in Brus and Saby (2016). It is a model-based estimator calculated using values retrieved from the semivariogram (De Gruijter et al. 2006). The semivariogram resembles the correlogram, but the autocorrelation $\hat{\rho}_h$ is replaced by the semivariance $\hat{\gamma}_h$. The semivariance of a pair of observations x_1 and x_2 is $\hat{\gamma}_{x_1,x_2} = \frac{1}{2}(x_1 - x_2)^2$ and the overall semivariance at a particular distance (or range) h is the mean semivariance of the pairs of observations separated by that particular distance (or in that lag):

$$\gamma_h = \frac{1}{2m} \sum_{i=1}^m (x_{i1} - x_{i2})^2$$

where m is the number of pairs $\{x_{i1}, x_{i2}\}$ separated by a distance falling within lag h . This empirical semivariance $\hat{\gamma}_h$ is considered to be an estimate of a theoretical semivariance γ_h , usually expressed as a function of h . The empirical semivariogram is a graph showing $\hat{\gamma}_h$ along the vertical axis and the corresponding separation distance h along the horizontal axis. An example is found in Fig. 7. The semivariogram was compiled for each land cover class, based on the available data for the entire population.

The variance estimator $VAR(\bar{x})$ MB for a particular land cover class and sampling distance was calculated as

$$VAR(\bar{x}) MB = \gamma_A - E_p[\gamma_{SY}]$$

where γ_A is the mean semivariance between two random points inside the study area, irrespective of separation distance, obtained by Monte Carlo simulation of a very large number of pairs of randomly located points (located inside the study area). The semivariance for each pair was retrieved from the semivariogram based on the distance between the points. The value γ_A was calculated as

$$\gamma_A = \frac{1}{M} \sum_{i=1}^M \hat{\gamma}_i$$

where M is the number of pairs in the simulation and $\hat{\gamma}_i$ is the semivariance for the distance separating the points in pair i . M has to be sufficiently large to allow γ_A to converge towards a stable value.

$E_p[\gamma_{SY}]$ is the “estimated p -expectation of the estimated mean semivariance within the sampling grid” (Brus and Saby 2016, p. 79):

$$E_p[\gamma_{SY}] = \frac{1}{k} \sum_{i=1}^k \frac{1}{n_i^2} \sum_{a=1}^{n_i} \sum_{b=1}^{n_i} \hat{\gamma}_{ab}$$

where $k = d^2$ is the number of partitions at sampling distance d and n_i is the number of sample points in partition i . $\hat{\gamma}_{ab}$ is the semivariance for the distance separating observations a and b . When the study area is irregular, as in this study, n_i will vary somewhat from one sample to another. Clearly, this approach requires that all the locations in the population are known, although the attribute value of most of the locations remains unknown.

3. Results

The exact (SYS) and mean estimated (SRS, ST4, V08 and MB) variance of \hat{x} for seven land cover classes at sampling interval 10 km are found in Table 2, summarizing the results from all the 100 partitions at this sampling interval. The table shows how SRS overestimate the true variance (SYS) while ST4 yield results fairly close to SYS. The results from MB are also close to SYS. Both ST4 and MB underestimated the variance for built-up and agricultural land. ST4 was in both cases the best estimate of the two. For the five other land cover classes, both ST4 and MB overestimated the variance.

Table 2

Exact (SYS) and mean estimated (SRS, ST4, V08 and MB) variance of \bar{x} for seven land cover classes at sampling interval 10 km, along with the (exact) spatial autocorrelation for this sampling interval.

Land cover class	VAR (\bar{x}) SYS	VAR (\bar{x}) SRS	VAR (\bar{x}) ST4	VAR (\bar{x}) V08	VAR (\bar{x}) MB	ρ_{10}
1 Built-up land	0.0050	0.0059	0.0041	0.0020	0.0034	0.1311
2 Agriculture	0.0213	0.0388	0.0235	0.0087	0.0180	0.2518
3 Forest	0.1217	0.3786	0.1737	0.0549	0.1813	0.4139
4 Open land	0.1440	0.4844	0.2148	0.0654	0.1835	0.4415
5 Mire	0.0186	0.0455	0.0262	0.0104	0.0236	0.2422
6 Snow/Ice	0.0062	0.0168	0.0096	0.0039	0.0124	0.2386
7 Water	0.0341	0.0567	0.0400	0.0229	0.0441	0.0685

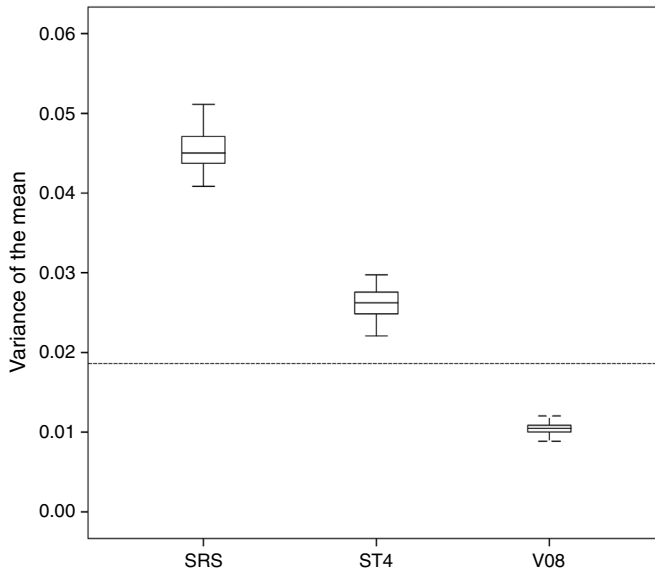


Fig. 3. Distribution of the variance of the mean coverage of Mire and peat bog, estimated with three methods (SRS, ST4 and V08) from a population divided into 100 partitions by systematic sampling on a 10 by 10 km frame. The exact variance for the mean of the 100 systematic samples ($VAR(\bar{x}) SYS$) is shown as a reference line at 0.0186.

ST4 gave the best estimate for three of these land cover classes, while MB was best for the remaining two classes. V08 underestimated the variance for all seven land cover classes.

A full comparison of the distribution of variance estimates for the 100 partitions at sampling interval 10 km for the land cover type Mire and peat bogs is shown in Fig. 3. The exact variance for the mean of the 100 systematic samples ($VAR(\bar{x}) SYS$) is 0.0186. This value is shown as a dotted horizontal line in the graph. The estimates using the SRS method return results in the range 0.04 to 0.05, well above the correct value. The ST4 method returns estimates in the range 0.02 to 0.03, slightly above the correct value. The V08 method returns estimates close to 0.01, which is below the correct value. MB is not included in the graph, since MB was calculated using the semivariogram for the entire population and not individual semivariograms for each partition. The single estimate of MB fell inside the range of ST4.

The variance estimates based on SRS, ST4 and V08 were standardized against SYS e.g.

$$VAR(x) SRS_{standardized} = 100 + (VAR(x) SRS - VAR(x) SYS) \times 100 / VAR(x) SYS$$

in order to facilitate comparison across the seven land cover classes. The result (Fig. 4) shows the distribution of these three estimators relative to SYS for all seven land cover types and with sampling distance 10 km. SYS is represented as the variance = 100 line. Each land cover class is represented with

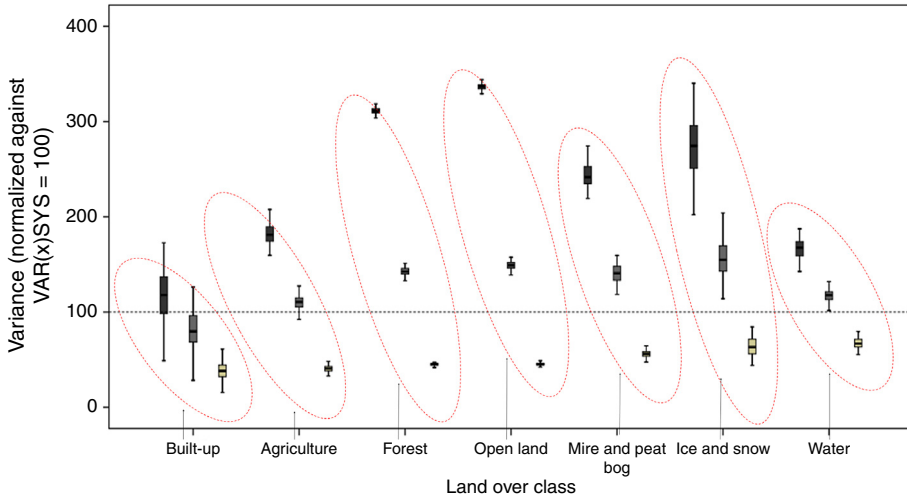


Fig. 4. Distribution of the variance of the mean coverage of seven different land cover types, each estimated with three methods (SRS, ST4 and V08) from a population divided into 100 partitions by systematic sampling on a 10 by 10 km frame. The estimates are standardized against the exact variance for the mean of the 100 systematic samples ($VAR(\bar{x})SYS$) set to 100. Each land cover type is encircled with a dashed line.

a triplet of bar graphs, encircled with a dashed line. Each triplet has the same general interpretation as the three bars in the bar graph in Fig. 3, explained above. The triplet representing Mire and peat bog is number five from the left and is identical to the bar graph in Fig. 3, except that the variances are standardized against $VAR(\bar{x})SYS = 100.0$.

Fig. 4 shows that V08 underestimated the variance for all seven land cover types, for this dataset and at sampling interval 10 km. ST4 underestimated the variance for built-up land and occasionally also for agricultural land. For built-up land, even SRS occasionally underestimated the variance from systematic sampling in this dataset and at the sampling interval 10 km. In most situations, though, ST4 slightly overestimated the variance, but less so than SRS. For six out of seven land cover classes, ST4 was also a more accurate estimate than V08. The exception was the land cover class Ice and snow where V08 was the best estimate in terms of approximation of the true variance. Still, V08 did also for this class lead to an underestimation of the true variance. MB was not included in Fig. 4.

The exact (SYS) and mean estimated (SRS, ST4, V08 and MB) variance of \hat{x} for the land cover class Mire and peat bogs for sampling intervals over the whole range from 2 to 20 km are found in Table 3. The table also includes the exact spatial autocorrelation for Mire and peat bogs at each sampling distance calculated from the entire dataset. SRS usually returned the highest estimates, generally overestimating the variance. V08 usually gave the lowest estimates, mostly underestimating the variance. ST4 and MB gave the best estimates. At sampling distances less than 5 km, ST4 and MB both overestimated the variance, and ST4 was the more precise of the two estimators. The pattern was inconsistent at sampling distances above 5 km. Here, MB usually gave slightly lower estimates than ST4 (but not always), but none of the two estimators were consistently better than the other.

(Notice that the column MPB in Table 3 refers to an alternative to V08 explained in the Discussion below. The column has been included in Table 3 in order to avoid an extra table in the paper.)

Part of the data from Table 3 is visualized in Fig. 5 where the exact variance (SYS) is drawn as a thick, solid line. The line shows how the empirically determined variance of systematic sampling is increasing with increasing sampling distance (and the samples correspondingly include a smaller part of the population). The naïve estimator, SRS, is systematically overestimating the variance of the systematic sample, and more so as the distance is increasing. SRS is seen as a dotted line marked with small squares well above the other lines in the graph. V08, on the other hand, systematically underestimates the variance. V08 is seen as a dotted line marked with small '+' symbols well below

Table 3

Exact (SYS) and mean estimated (SRS, ST4, V08 and MB) variance of \bar{x} for the land cover class Mire and peat bogs at sampling intervals ranging from 2 to 20 km, along with the (exact) spatial autocorrelation for the sampling intervals. There is no sampling at 1 km intervals since the sampling units are 1 km² plots. MPB is a variance estimate using the same general method as V08, but with a correction factor calculated from the empirical correlogram.

Sampling interval (km)	VAR(\bar{x}) SYS	VAR(\bar{x}) SRS	VAR(\bar{x}) ST4	VAR(\bar{x}) V08	VAR(\bar{x}) MB	VAR(\bar{x}) MPB	ρ_d
0							1.000
1							0.707
2	0.0002	0.0014	0.0005	0.0001	0.0021	0.0000	0.526
3	0.0006	0.0037	0.0016	0.0005	0.0044	0.0002	0.440
4	0.0028	0.0069	0.0034	0.0011	0.0040	0.0009	0.386
5	0.0051	0.0110	0.0057	0.0019	0.0067	0.0020	0.346
6	0.0109	0.0161	0.0085	0.0030	0.0094	0.0038	0.315
7	0.0091	0.0221	0.0121	0.0044	0.0108	0.0063	0.291
8	0.0138	0.0290	0.0161	0.0062	0.0130	0.0095	0.270
9	0.0158	0.0368	0.0209	0.0082	0.0158	0.0135	0.254
10	0.0186	0.0455	0.0262	0.0104	0.0236	0.0184	0.242
11	0.0326	0.0552	0.0326	0.0130	0.0242	0.0239	0.232
12	0.0420	0.0657	0.0390	0.0159	0.0343	0.0302	0.224
13	0.0421	0.0772	0.0468	0.0191	0.0358	0.0373	0.216
14	0.0444	0.0897	0.0533	0.0227	0.0497	0.0453	0.208
15	0.0607	0.1030	0.0605	0.0265	0.0493	0.0542	0.202
16	0.0687	0.1173	0.0702	0.0308	0.0602	0.0640	0.196
17	0.0746	0.1325	0.0791	0.0351	0.0702	0.0747	0.193
18	0.0920	0.1485	0.0899	0.0399	0.0780	0.0856	0.189
19	0.1368	0.1656	0.1028	0.0450	0.0979	0.0977	0.185
20	0.1257	0.1835	0.1076	0.0504	0.1225	0.1107	0.182

the other lines in the graph. ST4 seems to represent a fairly good approximation to SYS over the whole range of different sampling distances. ST4 is represented by the dotted line close to the SYS line, and marked with small triangles. (Notice that Fig. 5 also includes a graphical line called MPB and marked with small circles. MPB refers to an alternative to V08 which is explained in the Discussion below. The graphics has been included in Fig. 5 in order to avoid an extra figure in the paper.)

The spatial autocorrelation reported in Tables 2 and 3 are exact figures calculated from the entire population. These will not be available in a sampling survey where only part of the population is known. The calculation of V08 relies on an estimate of the spatial autocorrelation calculated from the observations included in the sample. All calculations of V08 in this study are done that way, using a spatial autocorrelation calculated from the sample itself. Clearly, this approach is using an estimate of the spatial autocorrelation which itself involves some uncertainty. This uncertainty is visualized in Fig. 6 where the distribution of the estimated spatial autocorrelation for Mire and peat bog at each sampling interval is shown as a boxplot. The figure illustrates how the uncertainty of the spatial autocorrelation estimate is increasing as the spatial autocorrelation itself is decreasing with increasing sampling distance.

4. Discussion

The efficiency and accuracy of a spatial sampling method and the associated statistical inference are determined by the properties of the population, the sampling strategy and the estimator that is used (Wang et al., 2013). The properties of the population are therefore of particular interest in spatial sampling and statistical inference. The properties of a population can be independent and identically distributed (i.i.d.), spatially autocorrelated or characterized by spatially stratified heterogeneity.

The data used in the present study are not iid. This is demonstrated by the autocorrelation coefficients (as in Tables 2 and 3) and the semivariograms (as in the example in Fig. 7).

Systematic sampling is expected to be more efficient than simple random sampling when autocorrelation is present. The data used in this study therefore represents a situation where systematic sampling is suitable. The relative advantage of the systematic sampling approach is related to the size (and probably also the range) of the autocorrelation effect.

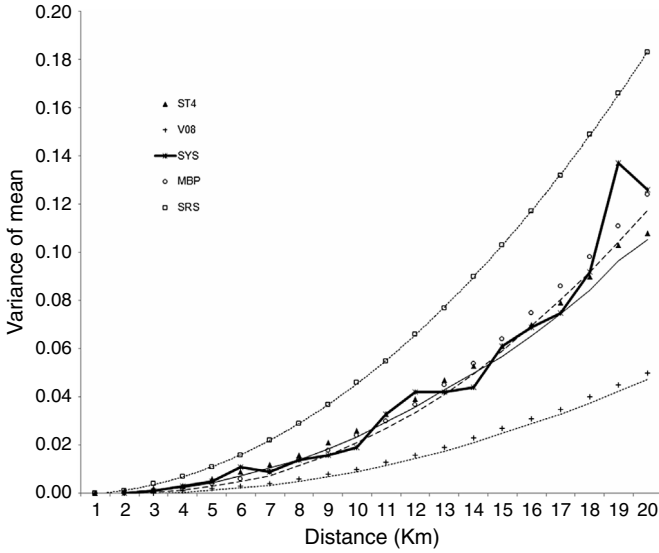


Fig. 5. Variance of the estimated mean coverage of Mire and peat bog from systematic sampling with sampling intervals in the range 2 to 20 km. The exact variance (SYS) is shown as a solid line, together with estimated variance using four different methods (STS, ST4, V08 and MPB) explained in the text.

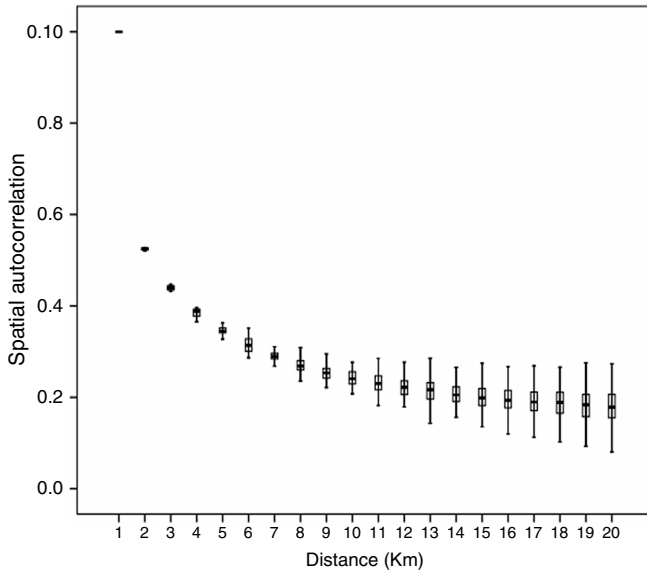


Fig. 6. Spatial autocorrelation for Mire and peat bog estimated using the available systematic samples at sampling intervals ranging from 2 to 20 km. Box plots are used to visualize the distribution of the estimates at each sampling distance. Spatial autocorrelation for neighboring samples (separation 1 km) is calculated from the entire population.

The semivariogram (an example is found in Fig. 7) for all seven land cover types in this study are rather complex. They exhibit signs of drift and periodicity. This impression is supported by maps showing the regional distribution of the land cover types (an example is found in Fig. 2). The explanation may be related to spatially stratified heterogeneity, since it is reasonable to expect that

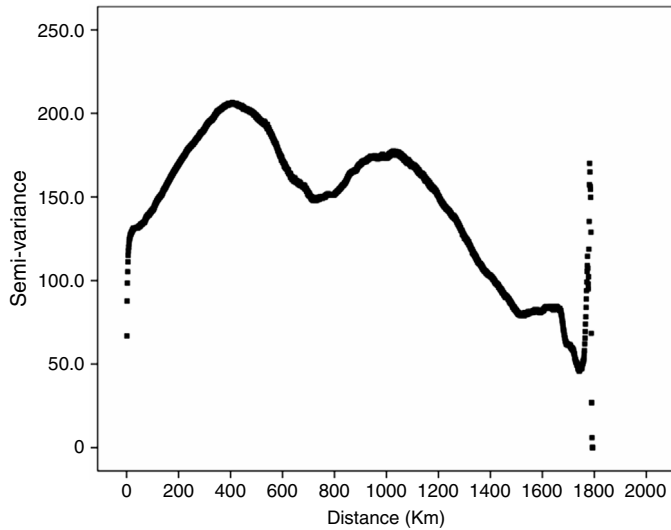


Fig. 7. Empirical semivariogram for the land cover class Mire and peat bog.

the distribution of land cover types is closely related to (e.g.) elevation above sea level, distance from the coast, latitude and climatic regions, resulting in spatially stratified heterogeneity.

Wang et al. (2016) have devised a method to measure spatially stratified heterogeneity, which could be used to gain more profound insight into the spatial properties of the data. The objective of the present study is, however, to compare variance estimation methods. In depth descriptions of the properties of the population falls outside the scope of the paper and is therefore not elaborated, but should still be considered when the results are interpreted.

The previous study by McGarvey et al. (2016) did not compare ST4 and V08, but found V08 to be a fairly precise estimator of the variance of a systematic sample when compared to a large number of other estimators. V08 did, however, not accomplish well in the present study. This is also in accordance with results reported by Brus and Saby (2016), who found that V08 (referred to as Moran in their paper) underestimated the variance.

The explanation for the divergence between results of the present study and the results reported by McGarvey et al. (2016) is probably related to the datasets used in the two studies. McGarvey et al. (2016) used synthetic datasets with strict control of how the population was generated. The present study is carried out with empirical data retrieved from a national geospatial database, entailing data with an irregular and erratic behavior.

Table 3 lists the true spatial autocorrelation at each sampling interval, calculated using the entire population. These figures will not be available in a sampling situation, where an estimated value for the spatial autocorrelation has to be computed from the sample alone. The result is seen in Fig. 6 where the distribution of the estimated spatial autocorrelation at each sampling interval are shown as boxplots. The estimated spatial autocorrelation is often different from the true spatial autocorrelation, and more likely to be so as the sampling interval is increasing. The correction factor arising from the spatial autocorrelation will vary accordingly: An estimated spatial autocorrelation of 0.3 results in a correction factor close to 0.20 while an estimated spatial autocorrelation of 0.2 results in a correction factor close to 0.26.

A second and possibly more serious source of differences between the results obtained with V08 here and in McGarvey et al. (2016) is the fact that the correction factor proposed by Cochran (1946) and used in V08 is based on the assumption that the spatial autocorrelation function has the form $\rho_d = e^{-\lambda d}$. As seen in Fig. 6, this is not the case here. The shape of the spatial autocorrelation function for Mire and peat bog in Norway is closer to $\rho_d = d^{-0.6}$. There is actually no reason to expect the

correction factor based on the explicit assumption of a particular exponential spatial autocorrelation function to work in situations where this assumption is violated. It is therefore reasonable to assume that V08 produced good results in the study carried out by [McGarvey et al. \(2016\)](#) – as well as in the older studies by [Osborne \(1942\)](#) and [Matèrn \(1947\)](#) – because the assumption about the spatial autocorrelation function was correct for the data used in those studies.

It is possible to construct an adjusted V08 by using a new correction factor based on the spatial autocorrelation function $\rho_d = d^{-0.6}$. Starting from the relationship $\frac{\text{VAR}(\bar{x})_{\text{SYS}}}{\text{VAR}(\bar{x})_{\text{SRS}}} = f(\rho_d)$, simple curve estimation using SPSS[®] shows that for Mire and peat bog $f(\rho_d) = 1.4e^{-4.07\rho_d}$ and an alternative to V08 for Mire and peat bog is thus $\text{VAR}(\bar{x}) \text{MPB} = \frac{1.4}{n} s_x^2 e^{-4.07\rho_d}$. The mean $\text{VAR}(\bar{x}) \text{MPB}$ for each sampling distance is listed in [Table 3](#) and shown graphically as a line marked with small circles in [Fig. 5](#). $\text{VAR}(\bar{x}) \text{MPB}$ slightly underestimates $\text{VAR}(\bar{x}) \text{SYS}$, but is fairly similar to $\text{VAR}(\bar{x}) \text{ST4}$ at small sampling distances (where the spatial autocorrelation is high) and slightly more conservative than $\text{VAR}(\bar{x}) \text{ST4}$ at larger sampling distances, where the latter tends to underestimate the variance. $\text{VAR}(\bar{x}) \text{MPB}$ always provides better estimates than $\text{VAR}(\bar{x}) \text{V08}$. The problem is, however, that $\text{VAR}(\bar{x}) \text{MPB}$ relies on detailed knowledge of the population and is only valid for this particular survey of Mire and peat bog in Norway.

The conclusion with respect to V08 is that this estimator, and similar estimates using the autocorrelation itself as a correction parameter, depends on the shape of the autocorrelation function. This function is unknown in most situations where a systematic sampling strategy is used. Without sufficient knowledge about the particular autocorrelation function for the dataset at hand, V08 cannot be used without risking underestimation of the true variance. Even when the autocorrelation function is known, it is not sufficient to calculate the autocorrelation value for the sampling interval and plug it into V08. The entire correction function has to be adjusted in order to fit the actual shape of the autocorrelation function.

The model-based estimator MB did well in this study, even with very complex semivariograms for all seven land cover classes. An example is the semivariogram for the land cover class Mire and peat bog shown in [Fig. 7](#). The semivariogram exhibits signs of autocorrelation, but also drift and periodicity. The function for distances h in the range 0–50 km is $\gamma(h) = e^{(4.9 - \frac{0.8}{h})}$ but other functions have to be fitted for other parts of the graph.

The number of random pairs that had to be drawn in order to stabilize the value of γ_A for use in the prediction of MB was very large. The requirement may be lower with a less complex semivariogram, but it is in any case necessary to ensure that the estimated γ_A is converging towards a stable value. MB may therefore be quite demanding with respect to computational power.

The study was carried out with a single semivariogram for each land cover class, compiled using the entire population. This approach does not document how the MB estimate will vary when the semivariogram is determined by using data from the sample instead of the entire population. [Figs. 3](#) and [4](#) show how the ST4 estimate varies between samples. The estimated MB was close to the mean estimated ST4, but the study does not document how much MB itself can vary. Further studies are recommended in order to shed light on this issue.

[Brus and Saby \(2016\)](#) found that MB could be approximated using post-stratification, with an estimator (STSI) resembling ST4. The difference between STSI and ST4 is that the post-stratification in ST4 is carried out with fixed segments of up to four sample units, while the post-stratification in STSI relies on k-means partitioning to divide the available observations into even smaller strata of two sample units each. The k-means partitioning is available in the R-package *spsosa* ([Walvoort et al., 2010](#)). The fact that MB can be approximated with a post-stratification estimator fairly similar to ST4 supports the impression that MB and ST4 are comparable and interchangeable as estimators of the variance in systematic samples.

5. Conclusion

This study has compared three estimators of the variance of parameters (e.g. mean value) determined using systematic random sampling: One estimator was based on post-stratification of the data (ST4), one was using a correction factor based on the observed spatial autocorrelation (V08) and

one was model-based using data retrieved from the semivariogram (MB). All three estimators have recently been tested and recommended in peer-reviewed and published papers (Aune-Lundberg and Strand, 2014; McGarvey et al., 2016; Brus and Saby, 2016). The contribution of the present study is that the estimators are compared using the same data.

In the present study, ST4 turned out to be fairly accurate or slightly conservative while V08 consistently underestimated the variance. The most likely explanation of the difference between the stratification approach (ST4) and V08 is that V08 relies on assumptions about the spatial autocorrelation function that is violated by the empirical data used in the experiment. The testbed was a real, national dataset retrieved from the Norwegian geospatial data infrastructure. The results do not disprove V08 as possibly the best estimator when the assumption about the spatial autocorrelation function having the form $\rho_d = e^{-\lambda d}$ is satisfied, but this assumption cannot be taken for granted. ST4 and MB are both better choices than V08 when the assumption about the shape of the spatial autocorrelation function is violated. The safe approach is therefore to use ST4 or MB.

The results obtained from MB and ST4 were quite similar. It is, however, not known how much MB will vary when the semivariogram is compiled from the sample instead of the population. A new study where MB is calculated using a semivariogram compiled from the sample itself, is therefore recommended in order to examine how sensitive MB is to inaccuracies in the autocorrelation model.

The conclusion of the current study is that ST4 and MB both can be recommended as variance estimators for systematic spatial samples. The main difference is that ST4 can be used when the spatial autocorrelation structure is unknown, while MB requires that a sufficiently accurate semivariogram is available.

Acknowledgment

The gridded datasets used in this study were compiled by Linda Aune-Lundberg as part of her Ph.D. project.

References

- Ambrosio, L., Iglesias, L., Marín, C., Del Monte, J.P., 2004. Evaluation of sampling methods and assessment of the sample size to estimate the weed seedbank in soil, taking into account spatial variability. *Weed Res.* 44, 224–236.
- Aubry, P., Debouzie, D., 2000. Geostatistical estimation variance for the spatial mean in two-dimensional systematic sampling. *Ecology* 81, 543–553.
- Aune-Lundberg, L., Strand, G.H., 2014. Comparison of variance estimation methods for use with two-dimensional systematic sampling of land use/land cover data. *Environ. Model. Softw.* 61, 87–97.
- Aune-Lundberg, L., Strand, G.H., 2017. Composition and spatial structure of dwarf shrub heath in Norway. *Nor. Geogr. Tidsskr. - Norw. J. Geogr.* 70, 1–11.
- Bellhouse, D.R., Sutradhar, B.C., 1988. Variance estimation for systematic sampling when autocorrelation is present. *Statistician* 37, 327–332.
- Brus, D.J., Saby, N.P.A., 2016. Approximating the variance of estimated means for systematic random sampling, illustrated with data of the French soil monitoring network. *Geoderma* 279, 77–86.
- Cochran, W.G., 1946. Relative accuracy of systematic and stratified random samples for a certain class of populations. *Ann. Math. Statist.* 17, 164–177.
- Cochran, W.G., 1977. *Sampling Techniques*, third ed. John Wiley & Sons, New York.
- De Gruijter, J., Brus, D.J., Bierkens, M.F., Knotters, M., 2006. *Sampling for Natural Resource Monitoring*. Springer Science & Business Media.
- D’Orazio, M., 2003. Estimating the variance of the sample mean in two-dimensional systematic sampling. *J. Agric. Biol. Environ. Stat.* 8, 280–295.
- Dramstad, W.E., Fjellstad, W.J., Strand, G.H., Mathiesen, H.F., Engan, G., Stokland, J.N., 2002. Development and implementation of the Norwegian monitoring programme for agricultural landscapes. *J. Environ. Manag.* 64, 49–63.
- Dunn, R., Harrison, A.R., 1993. Two-dimensional systematic sampling of land use. *Appl. Stat.* 42, 585–601.
- Eurostat, 2003. *The Lucas survey. European statisticians monitor territory*. Office for Official Publications of the European Communities, Eurostat, Luxembourg.
- Flores, L.A., Martinez, L.I., Ferrer, C.M., 2003. Systematic sample design for the estimation of spatial means. *Econometrics* 14, 45–61.
- Gallego, J., Delincé, J., 2010. The European land use and cover area-frame statistical survey. In: Benedetti, R., Bee, M., Espa, G., Piersimoni, F. (Eds.), *Agricultural Survey Methods*. John Wiley & Sons, Chichester, pp. 149–168.
- Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659–1673.
- Madow, W.G., Madow, L.H., 1944. On the theory of systematic sampling, I. *Ann. Math. Statist.* 15, 1–24.

- Martino, L., Fritz, M., 2008. New insight into land cover and land use in Europe. Land use/cover area frame statistical survey: Methodology and tools. *Statistics in focus* 33, Eurostat, Luxemburg.
- Matérn, B., 1947. Methods of estimation the accuracy of line and sample plot surveys. *Meddelanden från Statens Skogsforskningsinstitut* 36.
- Matérn, B., 1960. Spatial Variation. Stochastic models and their application to some problems in forest surveys and other sampling investigations. *Meddelanden från Statens Skogs- forskningsinstitut* 49, 1–155.
- McGarvey, R., Burch, P., Matthews, J.M., 2016. Precision of systematic and random sampling in partitioned populations: habitat patches and aggregating organisms. *Ecol. Appl.* 26, 233–248. <http://dx.doi.org/10.1890/14-1973.1>.
- McRoberts, R.E., Vibrans, A.C., Sannier, C., Nasset, E., Hansen, M.C., Walters, B.F., Lingner, D.V., 2016. Methods for evaluating the utilities of local and global maps for increasing the precision of estimates of subtropical forest area. *Can. J. Forest Res.* 46, 924–932.
- Milne, A., 1959. The centric systematic area-sample treated as a random sample. *Biometrics* 15, 270–297.
- Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A., Bellamy, P.H., Stephens, M., Kibblewhite, M.G., 2008. Soil monitoring in Europe: a review of existing systems and requirements for harmonisation. *Sci. Total Environ.* 391, 1–12.
- Opsomer, J.D., Francisco-Fernández, M., Li, X., 2012. Model-based non-parametric variance estimation for systematic sampling. *Scand. J. Stat.* 39, 528–542.
- Osborne, J.G., 1942. Sampling errors of systematic and random surveys of cover-type areas. *J. Amer. Statist. Assoc.* 37, 256–264.
- Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B., Inghe, O., 2011. National inventory of landscapes in Sweden (NILS) - scope, design, and experiences from establishing a multiscale biodiversity monitoring system. *Environ. Monit. Assess.* 173, 579–595.
- Strand, G.-H., 2013. The Norwegian area frame survey of land cover and outfield land resources. *Nor. Geogr. Tidsskr. - Norw. J. Geogr.* 67, 24–35.
- Strand, G.-H., Bloch, V.V.H., 2009. Statistical grids for Norway. Documentation of national grids for analysis and visualization of spatial data in Norway. *Statistics Norway*, Oslo.
- Thompson, S.K., 2002. *Sampling*, second ed. John Wiley & Sons, New York.
- Tomppo, E., Tuomainen, T., 2010. In: Tomppo, E., Gschwanter, T., Lawrence, M., McRoberts, R. (Eds.), *National Forest Inventories, Pathways for Common Reporting*. Springer, Finland, pp. 185–206.
- Tomter, S.M., Hysten, G., Nilsen, J.E., 2010. In: Tomppo, E., Gschwanter, T., Lawrence, M., McRoberts, R. (Eds.), *National Forest Inventories, Pathways for Common Reporting*. Springer, Norway, pp. 411–424.
- Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by *k*-means. *Comput. Geosci.* 36, 1261–1267.
- Wang, J.-F., Jiang, C.-S., Hu, M.-G., Cao, Z.-D., Guo, Y.-S., Li, L.-F., Liu, T.-J., Meng, B., 2013. Design-based spatial sampling: Theory and implementation. *Environ. Model. Softw.* 40, 280–288.
- Wang, J.-F., Stein, A., Gao, B.-B., Ge, Y., 2012. A review of spatial sampling. *Spat. Stat.* 2, 1–14.
- Wang, J.-F., Zhang, T.-L., Fu, B.-J., 2016. A measure of spatial stratified heterogeneity. *Ecol. Indic.* 67, 250–256.
- Wolter, K.M., 1984. An investigation of some estimators of variance for systematic sampling. *J. Amer. Statist. Assoc.* 79, 781–790.
- Wolter, K.M., 2007. Variance estimation for systematic sampling. In: Wolter, K.M. (Ed.), *Introduction to Variance Estimation*. Springer, New York, pp. 298–353.