International Journal of Statistics and Probability; Vol. 4, No. 4; 2015
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

A SAS Macro for Adaptive Spatial Sampling

Alan Ricardo da Silva[1] & Iracema Veiga Madeira Mauriz[1]

[1] Departamento de Estatistica, Universidade de Brasilia, Brazil

Correspondence: Alan Ricardo da Silva, Departamento de Estatistica, Universidade de Brasilia, Brazil. E-mail: alansilva@unb.br

Received: August 1, 2015 Accepted: August 25, 2015 Online Published: September 23, 2015

doi:10.5539/ijsp.v4n4p20 URL: http://dx.doi.org/10.5539/ijsp.v4n4p20

Abstract

This paper presents a SAS macro to estimate adaptive spatial sampling, which has been used to survey rare species. This technique is computationally difficult because of use of algorithms with GIS features such the creation of a grid, points inside polygons and contiguity. The results indicates that the SAS macro that was developed was capable of incorporating these GIS features, as well as estimating the parameters of the adaptive spatial sampling.

Keywords: adaptive sampling, spatial sampling, grid, SAS

1. Introduction

The purpose of sampling is to obtain information based on the results of a sample. According to Cochran (1977), sampling theory was developed to achieve the most efficient sampling, that is, to produce more accurate estimates with the lowest possible cost. Thus, the basic problem of any sampling procedure is to obtain reliable estimates of some characteristic of the population of interest, based on only part of this population.

A procedure that has been studied and tested in surveys of populations of rare species that display aggregated pattern distribution is adaptive sampling. In this kind of sampling, the selection of sampling units depends on observations made during the survey because if a criterion is met, the close sample is added to the initial sample. Thus, this type of sampling has advantages such as more extensive use of the sample and greater sampling intensity depending on the observations made during the survey; in addition, it can help find the local maximum (Thompson and Seber, 1996).

The adaptive sampling literature includes the following: algorithms that address the effects of mutations on the properties of folding RNA, the purpose of which is to decipher the principles of conduction and molecular evolution for the design of new molecules, in other words, these algorithms are for unbiased adaptive sampling that allows RNAmutants to sample regions of the mutational landscape that have not been fully addressed by previous techniques (Waldispühl and Ponty, 2011); designs for clinical study that focus on adaptation projects for two stage sample size re-estimation (Chang, 2008, 2009); mining applications that have a large amount of data, where random sampling may not be applicable due to the diffculty of determining an appropriate sample size (Domingo et al., 2002); and, finally, cases where the available algorithms for mining information on a large database are prohibitive due to computational constraints (time and memory) (Satyanarayana and Davidson, 2005).

In the case presented by Thompson (1990), another kind of adaptive sampling was considered, where the spatial distribution of aggregate data influences the formation of the sample selection of the data into clusters. This sample design, in which the procedure for the selection of units can be added to the initial sample based on an area and its spatial distribution, will be referred to from now on as adaptive spatial sampling.

Thus, adaptive spatial sampling provides a viable solution to the longstanding problem of estimating the abundance of rare populations and it has gained rapid acceptance in the natural and social sciences (Seber, 1986; Ramsey and Seber, 1992; Brown, 1994, 1996; Khan and Muttlak, 2002; Stein and Ettema, 2003; Sengupta and Sengupta, 2011; Jain and Chang, 2004; Thompson, 2011; Yu et al., 2012). However, adaptive procedures are more complicated to design and analyze, and computational implementations are few as a results of the complexity of the algorithms for spatial analysis (Thompson, 2011).

This implementation requires at least three steps: the development of a computational design for a regular grid; the selection of specific areas of the grid to identify which part of the grid the data are in; and identifying the neighbors of the selected areas: upper, lower, right and left. Thus, the objective of this work is to implement computationally adaptive spatial sampling in SAS software.

2. A Basic Outline of Adaptive Spatial Sampling

It has been observed that adaptive spatial sampling is performed in cluster sampling, because the data to be analyzed are divided into distinct subpopulations and have the geographical coordinates of a given area. Thus, the process of adaptive spatial sampling involves selecting certain areas, recording their geographical coordinates and defining the areas in which the data are located. Usually, these data are grouped in a particular area as shown in Figure 1(a), which characterize the clusters.

The first step is to draw a grid based on the geographical coordinates of the area to be analyzed, as shown in Figure 1(b). Then, some areas of the grid are selected by Simple Random Sampling (SRS) and it is determined whether there are data points within these areas, as shown in the blue areas of Figure 1(c). Next, the neighbors of these selected units are identified successively - top, bottom, right and left - until the selection criterion of the adaptive spatial sampling is exhausted. This adaptive spatial sampling is represented by one population sample (n) of the regular grid, as in Figure 1(d).

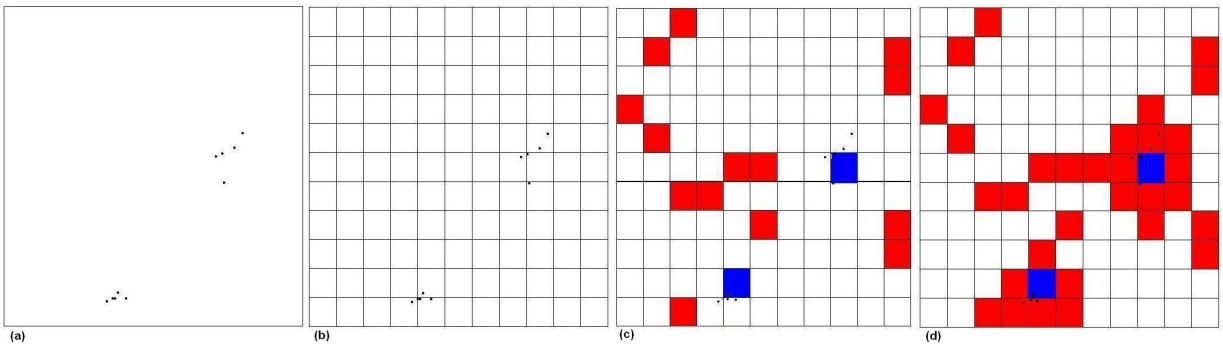

Figure 1. Steps of adaptive spatial cluster sampling

In summary, adaptive cluster sampling or simply adaptive spatial sampling refers to designs in which an initial set of units is selected by some probability sampling procedure and, whenever the variable of interest of a selected unit satisfies a given criterion, additional units in the neighborhood of that unit are added to the sample (Thompson, 1990). In the models considered in this paper, the initial sample can be selected by Simple Random Sampling with replacement SRS_R or without replacement SRS_{WR}.

2.1 Estimators

Classical estimators for the population mean are biased under an adaptive sampling design, in contrast with SRS. In this section two unbiased estimators for the population mean under an adaptive spatial sampling design will be addressed.

2.1.1 Estimators Using Initial Intersection Probabilities

This section shows an estimator based on a modification of a Horvitz-Thompson estimator (Thompson, 1990) and it is compared to the sample mean of the initial sample, given by

$$\bar{y} = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i \tag{1}$$

where y_i is the variable in study of the unit i and n_1 is the initial sample size.

When an initial sample n_1 of units is selected by a SRS_{WR}, these units in the first sample are distinguished not as a result of replacement. However, the data itself may contain repeated observations if more than one unit in the cluster is selected in the initial sample. The unit i will be included in the final sample if any unit of A_i (including i itself), where A_i is a neighborhood of the point i, is selected as part of the initial sample, or if any unit of a network of which unit i is an edge unit is selected, where an edge unit is a neighborhood of the point i but without sample points.

Let m_i denote the number of units in A_i; N is the population size; and a_i, the total number of units in the network (neighbors of the selected grid: upper, lower, right and left), of which unit i is an edge unit. Note that if the unit

satisfies the criterion C, i.e., some data point is found inside the selected grid, then $a_i = 0$, but if the unit i does not satisfy this condition, then $m_i = 1$. The selection probability of unit i in either n_1 observations is $p_i = \frac{m_i + a_i}{N}$. The probability that unit i is included in the sample is given by (Thompson, 1990):

$$\Pi_i = P(I_i = 1) = 1 - \left[\binom{N - m_i - a_i}{n_1} \Big/ \binom{N}{n_1} \right] \tag{2}$$

When the selection of the initial sample is taken by SRS_R, repeated observations in the data can occur either because of possible repeated selections in the initial sample or the initial selection of more than one unit in the cluster. In this sample design, selection probability of unit i in either n_1 observation is $p_i = \frac{m_i + a_i}{N}$, and the probability of inclusion is given by (Thompson, 1990):

$$\alpha_i = 1 - (1 - p_i)^{n_1} = 1 - \left(1 - \frac{m_i + a_i}{N}\right)^n \tag{3}$$

If the values of Π_i are known for all sample units, one can use the Horvitz-Thompson estimator given by $\widehat{\mu}_{HT} = \frac{1}{N} \sum_{i=1}^{n} \frac{y_i}{\Pi_i}$. However, although the values of m_i in Equation (2) for all units in the sample are known, only a few values of a_i are known. This means that i is a unit of edge somewhere in a cluster belonging to the sample, and thus, all clusters that this unit is related to do not need to be sampled. Thus, the value of a_i is unknown. To solve this problem, (Thompson, 1990) adopted the practice of dropping the value of a_i in Equation (2) and considering only the partial inclusion probability. Thus,

$$\Pi'_i = 1 - \left[\binom{N - m_i}{n_1} \Big/ \binom{N}{n_1} \right] \tag{4}$$

This probability Π'_i is now considered for n_1 networks instead of n_1 clusters and can be understood as the probability of the sample initial intercept A_i, the network for the unit i, to be used in the estimator. Thus, one obtains an unbiased estimator of the population mean based on the initial intersection probabilities as the following:

$$\widehat{\mu} = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i I'_i}{\Pi'_i} \tag{5}$$

where I'_i takes the value 1 (with probability Π'_i) if the initial sample intersects A_i, and 0 otherwise. In addition

$$\widehat{\mu}_{HT} = \frac{1}{N} \sum_{i=1}^{n} \frac{y_i}{\Pi_i} = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i I_i}{\Pi_i} \tag{6}$$

where $y_1 \ldots y_n$ represent the n distinct values of units in the final sample and I_i has the value 1 when the unit is included in the sample and 0 otherwise.

Using the properties of mathematical expectation it turns out that the estimator of Equation (5) is unbiased,

$$E[\widehat{\mu}] = E\left[\frac{1}{N} \sum_{i=1}^{N} \frac{Y_i I'_i}{\Pi'_i} \right] = \frac{1}{N} \sum_{i=1}^{N} \frac{Y_i E(I'_i)}{\Pi'_i} = \frac{1}{N} \sum_{i=1}^{N} \frac{Y_i \Pi'_i}{\Pi'_i} = \frac{1}{N} \sum_{i=1}^{N} Y_i = \mu \tag{7}$$

The classical estimator of the population mean under adaptive spatial sampling design is a biased estimator as follows:

$$E[\overline{y}] = E\left[\frac{1}{n_1} \sum_{i=1}^{n_1} y_i \right] = \frac{E\left(\sum_{i=1}^{N} y_i I'_i\right)}{n_1} = \frac{\sum_{i=1}^{N} Y_i E(I'_i)}{n_1} = \frac{\Pi'_i N \mu}{n_1} \neq \mu \tag{8}$$

To facilitate the analysis of Equation (5) it is more convenient to rewrite it in terms of distinct networks because the probability of intersection Π'_i is the same (also called α_k) for each unit i in the kth network. Thus,

$$\alpha_k = 1 - \left[\binom{N - x_k}{n_1} \Big/ \binom{N}{n_1} \right] \tag{9}$$

Similar to the equations of the probability of inclusion and as p_{jk} is the probability that kth and jth networks do not intersect, so

$$p_{jk} = P(J_j \neq 1 \cap J_k \neq 1) = \binom{N - x_j - x_k}{n_1} \Big/ \binom{N}{n_1} \tag{10}$$

where x_j is the number of units in the k-th network and J_k is the initial sample intersect of the kth network and takes the value 1 (with probability α_k) and 0 otherwise.

Using Equations (9) and (10) we obtain α_{jk} (the probability of intersection of the kth and jth networks) as

$$\begin{aligned}
\alpha_{jk} &= \alpha_j + \alpha_k - (1 - p_{jk}) \\
&= 1 - \left[\binom{N - x_j}{n_1}\right] \Big/ \binom{N}{n_1} + 1 - \left[\binom{N - x_k}{n_1} \Big/ \binom{N}{n_1}\right] - \left[1 - \binom{N - x_j - x_k}{n_1} \Big/ \binom{N}{n_1}\right] \\
&= 1 - \frac{\binom{N-x_j}{n_1}}{\binom{N}{n_1}} + 1 - \frac{\binom{N-x_k}{n_1}}{\binom{N}{n_1}} - 1 + \frac{\binom{N-x_{j-x_k}}{n_1}}{\binom{N}{n_1}} \\
&= 1 - \left[\binom{N - x_j}{n_1} + \binom{N - x_k}{n_1} - \binom{N - x_j - x_k}{n_1}\right] \Big/ \binom{N}{n_1}
\end{aligned} \tag{11}$$

Therewith,

$$\widehat{\mu} = \frac{1}{N} \sum_{k=1}^{K} \frac{y_k^* J_k'}{\alpha_k} = \frac{1}{N} \sum_{k=1}^{K} \frac{y_k^*}{\alpha_k} \tag{12}$$

where y_k^* is the sum of the y-values for kth network, K is the total number of distinct networks in the population, and k is the number of distinct networks in the sample.

Let $z_k = y_k^*/\alpha_k$, $y_k^* = \sum_{i=1}^{N} y_i\, \Pi_K = \alpha_k$ and Π_{jk}. From the properties of mathematical expectation, variance, covariance and the definitions above, one can obtain the expected value and the variance of Equation (5) by the following:

$$E[\widehat{\mu}] = \frac{1}{N} \sum_{k=1}^{K} z_k E(J_k) = \frac{1}{N} \sum_{k=1}^{K} z_k \alpha_k = \frac{1}{N} \sum_{k=1}^{K} y_k^* = \frac{1}{N} \sum_{i=1}^{N} y_i^* = \overline{y} = \frac{\tau}{N} = \mu \tag{13}$$

$$\begin{aligned}
var[\widehat{\mu}] &= var\left[\frac{1}{N} \sum_{k=1}^{K} z_k J_k\right] = \frac{1}{N^2} \left[\sum_{k=1}^{K} z_k J_k + \sum_{j=1}^{K} \sum_{j \neq k} cov\left(z_j J_j z_k J_k\right)\right] \\
&= \frac{1}{N^2} \left[\sum_{j=1}^{K} z_j^2 \Pi_j (1 - \Pi_k) + \sum_{j=1}^{K} \sum_{j \neq k} z_j z_k \Pi_{jk} - \Pi_i \Pi_k\right] \\
&= \frac{1}{N^2} \left[\sum_{j=1}^{K} \sum_{k=1}^{K} z_j z_k (\Pi_{jk} - \Pi_j \Pi_k)\right] \\
&= \frac{1}{N^2} \left[\sum_{j=1}^{K} \sum_{k=1}^{K} y_j^* y_k^* \left(\frac{\alpha_{jk} - \alpha_j \alpha_k}{\alpha_j \alpha_k}\right)\right]
\end{aligned} \tag{14}$$

and an unbiased estimator of the variance of Equation (14) is:

$$\begin{aligned}
\widehat{var}[\widehat{\mu}] &= \sum_{j=1}^{K}\sum_{k=1}^{K} z_j z_k J_j J_k \left(\frac{\Pi_{jk} - \Pi_j \Pi_k}{\Pi_{ij}}\right) \\
&= \frac{1}{N^2}\left[\sum_{j=1}^{K}\sum_{k=1}^{K} y_j^* y_k^* \left(\frac{\alpha_{jk} - \alpha_j \alpha_k}{\alpha_{jk}\alpha_j\alpha_k}\right) J_j J_k\right] \\
&= \frac{1}{N^2}\left[\sum_{j=1}^{K}\sum_{k=1}^{K} y_j^* y_k^* \left(\frac{\alpha_{jk}}{\alpha_{jk}\alpha_j\alpha_k} - \frac{1}{\alpha_{jk}}\right)\right] \\
&= \frac{1}{N^2}\left[\sum_{j=1}^{K}\sum_{k=1}^{K} \frac{y_j^* y_k^*}{\alpha_{jk}} \left(\frac{\alpha_{jk}}{\alpha_j\alpha_k} - 1\right)\right] \qquad (15)
\end{aligned}$$

Another known estimator for adaptive spatial cluster sampling is one that uses the expected number of initial intersection as follows.

2.1.2 Estimators Using the Expected Number of Initial Intersection

The estimator given by Equation (5) can be rewritten as:

$$\widetilde{\mu} = \frac{1}{N}\sum_{i=1}^{N} y_i \frac{f_i}{E[f_i]} \qquad (16)$$

where f_i represents the number of units in the initial sample that fall in network A_i, which includes the unit i; N is the number of regular grids. If during the estimation process the units edge of clusters is ignored, f_i would be interpreted as the number of times the ith unit of the final sample appears in the estimator. Then one realizes that $f_i = 0$ if no units in the initial sample intersect A_i.

The estimator in Equation (16) is unbiased because

$$E\left[\widetilde{\mu}\right] = E\left[\frac{1}{N}\sum_{i=1}^{N} y_i \frac{f_i}{E[f_i]}\right] = \frac{1}{N}\sum_{i=1}^{N} E(y_i)\frac{E[f_i]}{E[f_i]} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \mu \qquad (17)$$

Because m_i is the number of units on the network to which i belongs, using the Horvitz-Thompson estimator, another unbiased estimator can be found: As f_i units are selected from m_i units in A_i, f_i follows a hypergeometric distribution with the following parameters: (N, m_i, n_1) (Thompson, 1991). Thus, $E[f_i] = \frac{n_1 m_i}{N}$ and substituting the expected value in Equation (16) we obtain the following:

$$\widetilde{\mu} = \frac{1}{N}\sum_{i=1}^{N} y_i \frac{f_i}{\frac{n_1 m_i}{N}} = \frac{N}{N}\sum_{i=1}^{N} \frac{y_i f_i}{n_1 m_i} = \frac{1}{n_1}\sum_{i=1}^{N} \frac{y_i f_i}{m_i} \qquad (18)$$

To find the variance of the estimator of Equation (18), the approach in terms of n_1 networks being connected is used, although this is not necessarily distinct. Because m_i has the same value for all units in A_i and w_i is the average of m_i observations A_i (Thompson, 1990), then

$$\widetilde{\mu} = \frac{1}{n_1}\sum_{i=1}^{n_1} \frac{1}{m_i}\sum_{j\epsilon A_i} y_j = \frac{1}{n_1}\sum_{i=1}^{n_1} w_i = \overline{w} \qquad (19)$$

Thus, $\widetilde{\mu}$ is the sample mean obtained by taking a selection of an SRS of size n_1 of a population of w_i values rather then y_i values. Because $w_i = \overline{v}_k$ is the same for each unit in the kth network, where $\overline{v}_k$ is the mean of the y-values in β_k, there are x_k units in the kth network and B_K is a set of units in the kth network; thus,

$$E(\widetilde{\mu}) = E(\overline{w}) = E\left[\frac{1}{N}\sum_{i=1}^{N} w_i\right] = \frac{1}{N}\sum_{k=1}^{K} x_k \overline{v}_k = \frac{1}{N}\sum_{k=1}^{K}\sum_{i\epsilon B_k} y_i = \mu \qquad (20)$$

From the equations of cluster sampling in two stages, we obtain an unbiased estimator of this variance:

$$var[\widetilde{\mu}] = var\left[\frac{1}{N}\sum_{i=1}^{N} w_i\right] = \frac{N-n_1}{Nn_1(N-1)}\sum_{i=1}^{N}(w_i-\mu)^2 = \frac{\sigma^2}{n_1}\left(1-\frac{n}{N}\right) \tag{21}$$

where $\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(w_i-\mu)^2$.

$$\widehat{var}[\widetilde{\mu}] = \frac{N-n_1}{Nn_1(n_1-1)}\sum_{i=1}^{n_1}(w_i-\widetilde{\mu})^2 \tag{22}$$

As in some populations a priori information may be known; i.e., where aggregations occur, one can use the technique to reduce the stratified sample variance estimators. To do this, we can use stratified adaptive spatial sampling, where the population is divided into strata and the number of units are sampled for each stratum that is used.

3. Stratified Adaptive Spatial Cluster Sampling

In the case of adaptive spatial sampling techniques, one must also know the geographic coordinates of the selected area. Thereafter, one stratifies the area using a priori information and draws the grid throughout the selected area (including stratified areas) through their respective locations as indicated in Figure 2 (a). After that, a sample is selected using *S RS*, as in Figure 2 (b). Then, the units of interest are identified, as in Figure 2 (c). Finally, one successively adds the neighbors of selected areas - upper, lower, right and left - until the selection criterion of the adaptive sample is exhausted, as in Figure 2 (d).

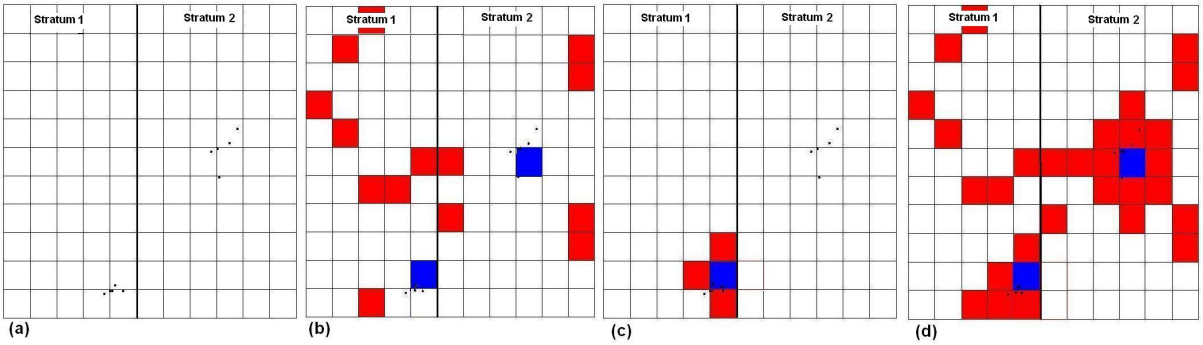

Figure 2. Steps of stratified adaptive spatial cluster sampling

In the estimators to be considered, the initial sample can also be selected by SRS_R or SRS_{WR}. The next section will show three unbiased estimators for this kind of sampling.

3.1 Estimators

Suppose that the population total of N units is partitioned into L stratum, with n_h units in the hth stratum ($h = 1, 2, \ldots, L$). Define unit (h, i) as the ith unit in the hth stratum with associated y-value y_{hi}. This process begins with a SRS of n_h units that is taken from stratum L, and we now define $n_0 = \sum_{h=1}^{L} n_h$ to be the initial total sample size. From this, the clusters begin to have neighbors added according to the condition set C (Thompson and Seber, 1996).

3.1.1 Estimators Using Initial Intersection Probabilities

Using the full adaptive sample, the first estimator we can consider is given by (5) based on the initial intersection probabilities, we obtain

$$\widehat{\mu}_{st} = \frac{1}{N}\sum_{k=1}^{K}\frac{y_k^* J_k}{\alpha_k} \tag{23}$$

where the K distinct networks are labeled $(1, 2, \ldots, k)$ without regard for stratum boundaries, J_k equals 1 (with probability α_k) if the initial sample size n_0 intersects network k, and 0 otherwise and, finally, y_k^* is the sum of the y-values for the network k.

To derive α_k it is necessary to consider the probabilities of intersecting network k with the initial samples in each strata. Therefore, we define x_{hk} as the number of units in stratum h that lie in network k. This number assumes the

value 0 if the network k lies totally outside stratum h. If the network straddles a boundary, we ignore the network units that lie outside *stratum h* in the definition of x_{hk}. Thus, with this definition of x_{hk}, we obtain

$$\alpha_k = 1 - \left[\prod_{h=1}^{L} \frac{\binom{N_h - x_{hk}}{n_h}}{\binom{N_h}{n_h}} \right] \tag{24}$$

The variance of the estimator of the unbiased average, defined as the probability of the initial sample intercede network in k and k', is obtained in the following way:

$$\alpha_{kk'} = 1 - (1 - \alpha_k) - (1 - \alpha_{k'}) + \left[\prod_{h=1}^{L} \frac{\binom{N_h - x_{hk} - x_{hk'}}{n_h}}{\binom{N_h}{n_h}} \right] \tag{25}$$

Because $\alpha_{kk} = \alpha_k$ and from the variance properties, it follows that

$$var[\widehat{\mu}_{st}] = \frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} y_k^* y_{k'}^* \left(\frac{\alpha_{kk'} - \alpha_k \alpha_{k'}}{\alpha_k \alpha_{k'}} \right) \tag{26}$$

and an unbiased estimator for this variance is given by

$$\widehat{var}[\widehat{\mu}_{st}] = \frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} y_k^* y_{k'}^* \left(\frac{\alpha_{kk'} - \alpha_k \alpha_{k'}}{\alpha_{kk'} \alpha_k \alpha_{k'}} \right) I_k I_{k'} \tag{27}$$

Another estimator for this kind of sampling is to use the expected number of the initial intersection, which will be explained in the next section.

3.1.2 Estimators Using the Expected Number of the Initial Intersection

Let A_{hi}, the network that contains the unit (h, i), u_{hi}, and A_{ghi} be part of A_{hi} stratum g. Suppose that f_{ghi} is the number of units from the initial sample in stratum g that fall in A_{ghi}, and let m_{ghi} be the number of units in A_{ghi}. Then, the number of units of an initial sample n_0 units is A_{hi} given by (Thompson and Seber, 1996),

$$f_{\cdot hi} = \sum_{g=1}^{L} f_{ghi} \tag{28}$$

From Equation (16) one obtains the estimator for the mean

$$\widetilde{\mu}_{st} = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} y_{hi} \frac{f_{\cdot hi}}{E[f_{\cdot hi}]} \tag{29}$$

As in Equation (7) and from the properties of expectation and variance, this estimator is unbiased.

As with f_i in Equation (16), f_{ghi} follows a hypergeometric distribution with parameters (N_g, m_{ghi}, n_g). Therefore, it is known that $E[f_{ghi}] = \frac{n_g m_{ghi}}{N_g}$ and $E[f_{\cdot hi}] = \sum_{i=1}^{L} \frac{n_g}{N_g} m_{ghi}$ (Thompson, 1991). Thus,

$$\widetilde{\mu}_{st} = \frac{1}{N} \sum_{h=1}^{L} y_{hi} \frac{f_{\cdot hi}}{\sum_{i=1}^{L} \frac{n_g}{N_g} m_{ghi}} = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} \left(y_{hi} \sum_{g=1}^{L} f_{ghi} / \sum_{g=1}^{L} \frac{n_g}{N_g} m_{ghi} \right) \tag{30}$$

where f_{ghi} represents the number of units in the initial sample that is at the intersection of stratum g with the network drive to which the unit u_{hi} belongs.

If there is a match to add the same neighbors, we obtain an estimator of the independent stratum combined with the estimator with weights, providing an estimator of the population mean as Equation $E[\overline{y}_{st}] = E\left(\sum_{h=1}^{L} \frac{N_h}{N} \overline{y}_h\right)$. This characteristic aggregation of equal neighbors generates a loss of efficiency, a more efficient system would allow groups to overlap the boundaries of the strata (Thompson, 1991).

So, to find the variance estimator of the mean, we use Equation (19) to rewrite $\widetilde{\mu}_{st}$ in terms of the weights of the sample means. For this, relate the observations to the intercept of the initial sample networks. Thus, the term $y_{hi}f_{.hi}$ means that A_{hi} is intersected f_{hi} times by the initial sample, so that $\widetilde{\mu}_{st}$ represents a weighted sum of all units in all the networks corresponding to the initial sample, with some networks being repeated. Because the weight $E[f_{.hi}]$ is the same for each unit in A_{hi}, we have

$$\widetilde{\mu}_{st} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{n_h}\frac{1}{E[f_{.hi}]}\sum_{(h',i')\in A_{hi}} y_{h'i'} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{n_h}\frac{Y_{hi}}{E[f_{.hi}]} \tag{31}$$

where Y_{hi} is the sum of yth observations in A_{hi}.

Let $\overline{w}_h = \sum_{i=1}^{n_h}\frac{w_{hi}}{n_h}$ and $w_{hi} = \frac{n_h Y_{hi}}{N_h E[f.hi]}$; then, another way to rewrite Equation (31) is

$$\widetilde{\mu}_{st} = \sum_{h=1}^{L}\frac{N_h}{N}\overline{w}_h = \frac{1}{N}\sum_{h=1}^{L}\frac{N_h}{n_h}\sum_{i=1}^{n_h} w_{hi} \tag{32}$$

where $w_{hi} = \frac{Y_{hi}}{\sum_g m_{ghi}}$; when $\frac{n_h}{N_h}$ have the same value for all strata. Thus, Equation (32) represents a stratified sample mean from a stratified random sampling without replacement, with the interest of w_{hi} variable.

So, the variance estimator for the mean is given by

$$var[\widetilde{\mu}_{st}] = \frac{1}{N^2}\sum_{h=1}^{L} N_h(N_h - n_h)\frac{\sigma_h^2}{n_h} \tag{33}$$

where σ_h^2 represents the stratum population variance, that is,

$$\sigma_h^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(w_{hi} - \overline{W}_h)^2 \tag{34}$$

where $\overline{W}_h = \frac{\sum_{i=1}^{n_h} w_{hi}}{n_h}$ is the stratum population mean.

An unbiased estimator of variance of the mean (33) can be obtained by replacing σ_h^2 by sample variance, $s_h^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(w_{hi} - \overline{w}_h)^2$.

3.1.3 Estimators that Ignore Units Added through Crossing Boundaries

According to Thompson (1991), the estimator that ignores units added through crossing stratum boundaries is given by the following:

$$\mu_{st}^{''} = \sum_{h=1}^{L}\frac{N_h}{N}\widetilde{\mu}_h = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h}\left(y_{hi}\sum_{g=1}^{L}\frac{N_g}{n_g}f_{ghi}/\sum_{g=1}^{L} m_{ghi}\right) \tag{35}$$

where $\widetilde{\mu}_h = \sum_{i=1}^{n_h}\frac{w_{hi}^{''}}{n_h}$ and $w_{hi}^{''}$ is the total of the y-values in the intersection of the stratum h with A_{hi} divided by the number of units in the intersection; i.e., this value represents the network mean for that part of the network A_{hi} in stratum h.

The mathematical expectation of $\mu_{st}^{''}$ is given by

$$E[\mu_{st}^{''}] = \sum_{h=1}^{L}\frac{N_h}{N}\mu_h = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{hi} = \mu \tag{36}$$

The variance $var[\mu_{st}^{''}]$ is given by

$$var[\widetilde{\mu}_{st}^{''}] = \frac{1}{N^2}\sum_{h=1}^{L} N_h(N_h - n_h)\frac{\sigma^2}{n_h} \tag{37}$$

where the stratum population variance is $\sigma_h^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(w_{hi}^{''} - \overline{W}_h)^2$ and the stratum population mean is $\overline{W}_h = \sum_{i=1}^{N_h}\frac{w_{hi}}{N_h}$. The sample estimate is given by (32) replacing σ_h^2 by $s_h^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(w_{hi} - \overline{w}_h)^2$.

4. SAS Macros

The SAS Macros basically use the IML Procedure and GMAP and SQL Procedures. The computational implementation of adaptive spatial sampling requires four steps: 1) development of the computational design of regular grids (Figure 3(a)); 2) selection of specific areas of the grid, that is, identifying a sample and determining in which part of the grid the data are located (Figure 3(b)); 3) identification of the neighbors of the selected areas: upper, lower, right and left (Figure 3(c)(d)); 4) calculation of the parameters.

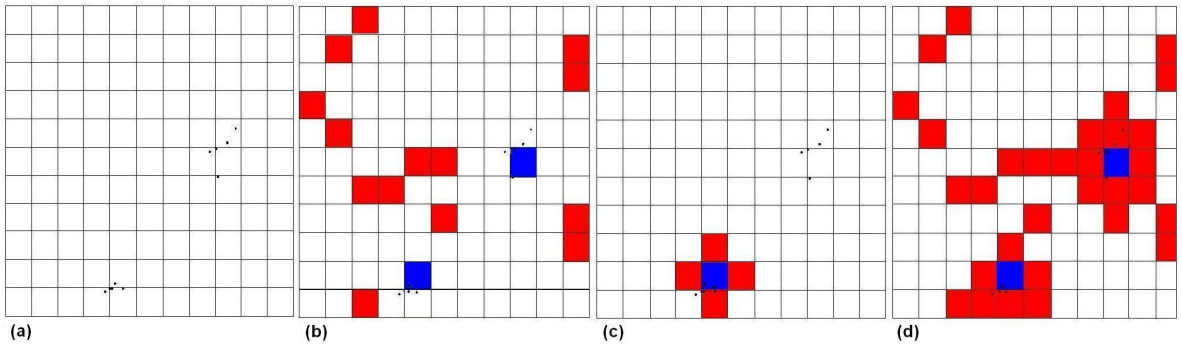

Figure 3. Steps of adaptive spatial sampling

4.1 Drawing a Regular Grid

To create regular grids it is necessary to create four points with coordinates entered clockwise (lines of Table 1) or counterclockwise.

Table 1. Coordinates of a square

Reference	Values	*Points*
1	**(Min, Min)**	(0,0)
2	**(Min, Max)**	(0,1)
3	**(Max, Max)**	(1,1)
4	**(Max, Min)**	(1,0)

In the case of a square, beginning with the clockwise points, i.e., the reference points in the following order: 1, 2, 3, 4, the polygon appears to be like Figure 4 (Square). If that order is not followed, the result is a distorted polygon, as shown in Figure 4 (Distorted Polygon).

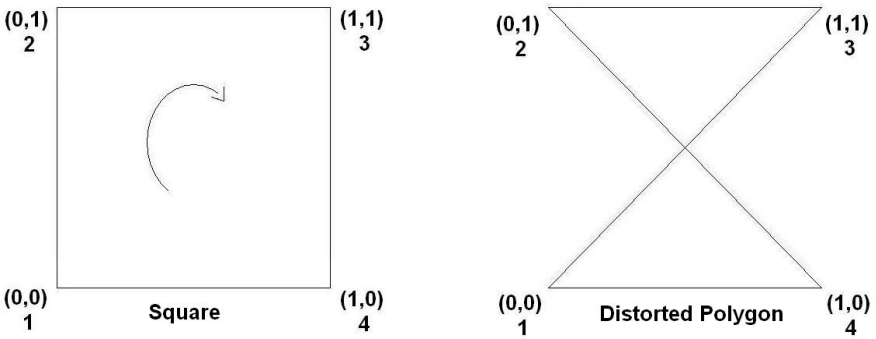

Figure 4. Square and distorted polygon.

As one wishes to draw a grid on a field of study, it is necessary to know the upper and lower boundaries of the region, i.e., the minimum and maximum coordinates of the *y* axis (latitude) and minimum and maximum coordinates of the *x* axis (longitude). The definition of the size of each polygon is given by **%grid** macro:

```
%grid(minx =, maxx =, miny =, maxy =, dim =, anno =, printN = YES );
```

where the parameters are: **MINX =** the minimum value of the **x** coordinate; **MAXX=** the maximum value of the **x** coordinate. Similarly, for the **y** coordinates we have: **MINY=** and **MAXY=**. Another parameter of this macro is the size of the square drawn given by **DIM=** (for instance, if **DIM=** 20, $20^2 = 400$ squares will be created). Finally, the last two parameters, **ANNO=** and **PRINTN = YES**, indicate the dataset with the location of samples and whether the numbering of each square will be printed using the command **YES**(Figure 5 (a)) or **NO** (Figure 5 (b)), respectively.

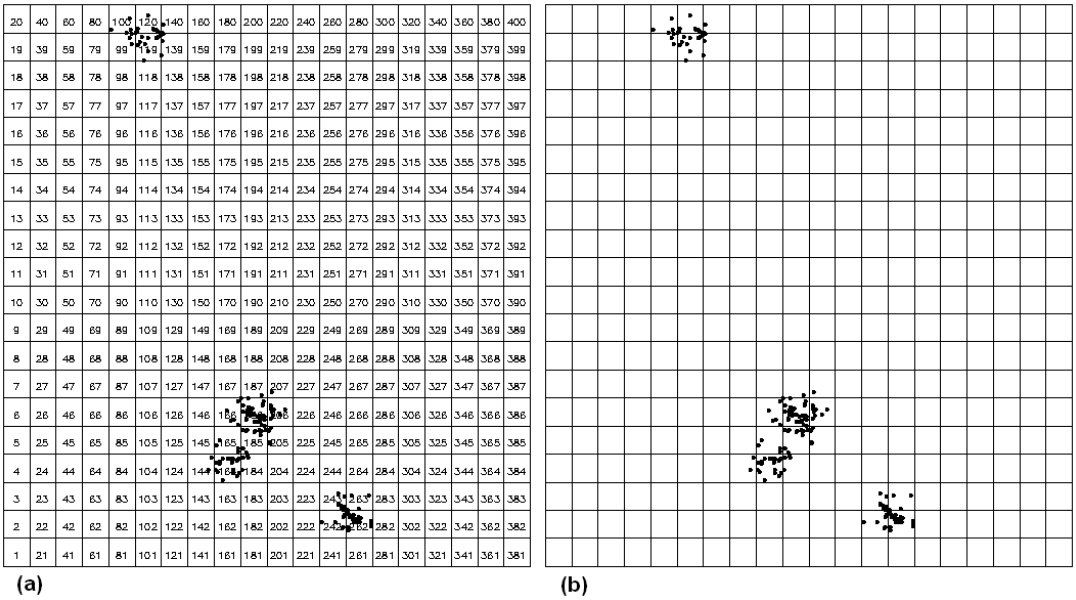

Figure 5. Regular grid for $N = 400$.

Next, we define the element, **id** to the coordinates of the square. Thus, the first square has the points $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, **id = 1** and so on. This is achieved by joining the table with the coordinates with the table set out below:

```
data id&dim;
 do id=1 to &dim*&dim;
  do i=1 to 4;
   output;
  end;
 end;
run;
```

This numbering starts from a unit numeric value and goes to the value of N to count the vertical direction, starting from left to right. In the case of Figure 5, it is found that the size of the square is $N = 20 \times 20 = 400$, varying the **id** from 1 to 400.

The next step is to make the selection of specific areas of the grid; i.e., the grids are drawn by SRS, and if there are samples inside a grid, it is selected by its neighbors.

4.2 Selection of Specific Areas of the Grid

The samples to be drawn in adaptive spatial sampling correspond to the polygons of the grid. This selection can be done by a generator corresponding to the number of grid squares of random numbers. In SAS, this can be done by PROC SURVEYSELECT, where a SRS is obtained with a seed value of size n given by the variable **SEED**. The parameter **OUT** indicates where the sample will be stored.

```
proc surveyselect data=&data sampsize=&n out=&saida seed=&seed noprint;
run;
```

The identification of the neighbors of the selected areas in the next section involves three concepts: check point inside the polygon; definition of the neighbors; and identification of neighboring polygons.

4.3 Identification of the Neighbors

The selection of neighbors is the trickiest and the most important part of the adaptive spatial sampling technique, as it is from the selected grids that the process of adapting the sample areas begins.

4.3.1 Checking if the Point is Inside the Polygon

The determination of points inside the polygon is shown in Figure 6. The main idea consists in making a radius from the selected point p to infinity in any direction and computing the amount of times that this line passed through the edges of the polygon. Thus, if the number of crossings is odd, the point is inside the polygon (Kunigami, 2010).

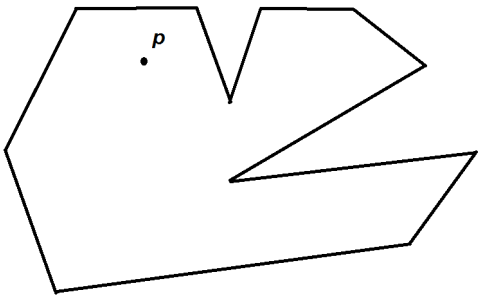

Figure 6. Point inside polygon.

The selection of these points within the regular grid is done by **%ginside** macro.

```
%ginside(map = , id = , where = , data = , out = );
```

where the elements of the macro are: **MAP =** dataset containing the coordinates of the study area; **ID=** name of the ID variable that defines the polygon; **WHERE =** selecting a specific point to check if it is inside the polygon; **DATA =** dataset containing the points to be checked if they are inside the polygon; **OUT =** dataset in which the points inside the polygons will be stored.

4.3.2 Definition of the Neighbors

The definition of the neighbors is as follows: The neighborhood is a set of squares that are added to the same sample if the grid satisfies the same condition of containing elements of interest in the selected square. We define the neighborhood of type ROOK as the polygons that share more than one point in common, in this case the square: up, down, right and left, as shown in Figure 3(c), and the neighborhood of type QUEEN as the polygons that share at least one point in common, i.e., the neighborhood ROOK adding the corners.

The **%neighborhood** macro is given by:

```
%neighborhood(id = , pt = , map = , anno = , out = , type = ROOK );
```

where the parameters are: **ID=** name of the ID variable that defines the polygon; **PT=** ID of the grid in which the neighbors will be defined; **MAP=** the dataset containing the coordinates of the study area; **ANNO=** the dataset containing the coordinates of the samples; **OUT=** the dataset in which the neighbors will be stored; **TYPE=** indicates that the pattern of selection of the neighbors is of type **ROOK** (default) or **QUEEN**.

4.3.3 Identifying Neighboring Polygons

The identification of neighboring polygons is generated by combining the identification of neighbors with the points inside the polygon; this is the final sample of adaptive spatial sampling.

Finally, given the base with the selected units and their respective score points, the next step is to compute the estimators presented in Section 2.

4.4 Estimators of Adaptive Spatial Sampling

In this section the formulas for the estimators of Section 2 were implemented. Thus, for adaptive spatial sampling we have implemented the estimator of the mean (**u1**), Equation (19), the estimator of the variance of the estimator

of the mean (**varu1**), Equation (21) the estimator of the total (**Totu1**) given by multiplying the number of grids by the estimator **u1**, i.e., **Totu1 = NN**× **u1**, and the estimator of the variance of **Totu1**, named **TotVaru1**.

The **%as** macro computes the estimators and automatically uses the macros presented previously.

```
%as(data = , n = , sample = , out = , strata = , seed = , map = , id = ,
    anno = , typen = ROOK , printN = YES );
```

where the parameters are the same as those presented earlier, and **n=** the sample size and **SAMPLE=** a dataset containing a predefined sample.

5. Illustration

5.1 An Example of Adaptive Spatial Cluster Sampling

Thompson (1990) presents an example of how adaptive spatial sampling works and compares the results obtained from the SRS estimator and from SRS with adaptive spatial sampling, which is given by changing the denominator of Equation (19) to the denominator of Equation (8). This example could represent a reserve of animals that are grouped (as herds of elephants) or deposits of minerals (such as gold, diamond, iron) spread over large areas.

Initially, a regular grid is drawn on the area to be sampled, and then, n units (squares) are selected by the SRS method. In this example, the initial sample consists of 10 units (total squares on the grid in red) selected in a total of $N = 400$ units (representing the total number of regular square grids, where each side has a length of 20, or $20 \times 20 = 400$), with a total of 190 points (Figure 7 (c)).

Selecting the neighbors (by the ROOK methodology) of the initial units containing at least one unit in the initial sample, we obtain the final sample, as shown in Figure 7(d). The upper unit has an element that intersects with the network $m_1 = 6$ units, containing a total of $y_1^* = 36$ units of interest. Another point in which there is unity within the polygon that intersects the network $m_2 = 11$ units and contains $y_2^* = 107$ units. For the other 8 units of the initial sample, the values are $y_i = 0$ and $m_i = 1$.

There are also 20 edge units that are not used in calculating the estimates; these are selected for adaptive selection, but they do not contain units of interest. In Figure 7(d) networks within the two groups that are described as being adaptively added are in the color red.

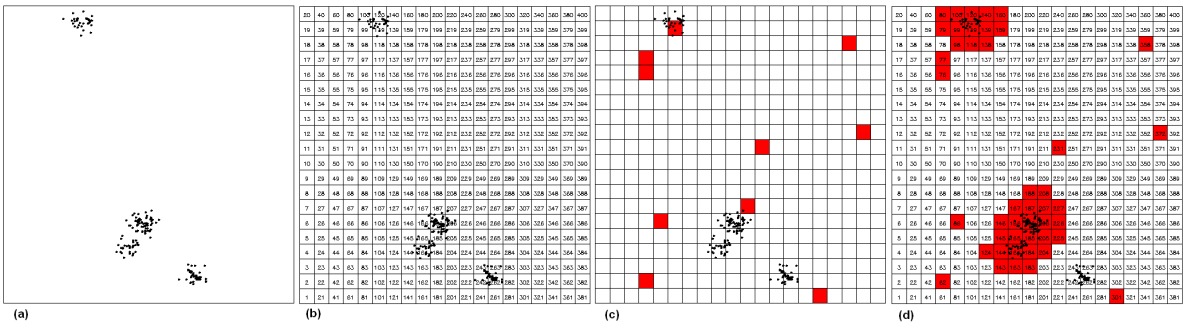

Figure 7. Example of adaptive cluster sampling.

For $w_1 = \frac{36}{6} = 6$ objects per unit, for $w_2 = \frac{107}{11} = 9.727$ and for the remaining $w_i = 0$, one can calculates the values for the mean estimators, $\widetilde{\mu}$, and from the total of the adaptive sampling:

$$\begin{aligned}
\widetilde{\mu} &= \frac{1}{10}\left[\frac{36}{6} + \frac{107}{11} + \left(\frac{0}{1}\right) + \ldots + \left(\frac{0}{1}\right)\right] = 1.573 \\
N\widetilde{\mu} &= 400 \times 1.573 = 629 \\
\widehat{var}[\widetilde{\mu}] &= \frac{(400-10)}{400(10)(10-1)}[(6-1.573)^2 + \ldots + (0-1.573)^2] = 1.147 \\
N^2\widehat{var}[\widetilde{\mu}] &= 400^2 \times 1.147 = 183,520
\end{aligned}$$

For SRS, where N is the total number of square areas selected and n is the number of selected squares of the initial

sample, one obtains the values for the estimators of the mean $\bar{y}$ and for the total $N\bar{y}$:

$$\begin{aligned} \bar{y} &= \frac{11+1}{10} = 1.2 \\ N\bar{y} &= 400 \cdot 1.2 = 480 \\ \widehat{var}[\bar{y}] &= 1.165 \\ N^2\widehat{var}[\bar{y}] &= 186,506 \end{aligned}$$

For the 45 units, which includes the 25 edge units of the final sample, the values for the estimators of the average $\bar{y}_{AD}$ for adaptive SRS are calculated; that is, the amount of SRS Equation in Equation (19) is used. Thus,

$$\begin{aligned} \bar{y}_{AD} &= \frac{143}{45} = 3.178 \\ N\bar{y}_{AD} &= 400 \times 3.178 = 1,271 \\ \widehat{var}[\bar{y}_{AD}] &= 1.004 \\ N^2\widehat{var}[\bar{y}_{AD}] &= 400^2 \cdot 1.004 = 160,687 \end{aligned}$$

Table 2 presents the estimates found and it appears that the variance, mean and total sampling of the adaptive SRS is the lowest compared to the others. However, their estimated average is higher, because there is a bias when using the estimator of SRS in this sample. Comparing the ratio of the variances $\bar{y}_{AD}$ and $\widetilde{\mu}$, we observe that there is a reduction of 13% in this value. Thus, given that it has a total of 190 points and the actual population mean is $\mu = \frac{190}{400} = 0.475$, adaptive spatial sampling in this case was very close to the SRS; however as will be seen later, adaptive spatial sampling varies much less when N varies.

The computational output of the program implemented in SAS software for this example is given in Figure 8 (the results are the same as in Thompson (1990)). Thus, there is the number of observations ($n = 10$), population size ($N = 400$), the estimators for the mean and total and their respective variances in the three cases analyzed: adaptive spatial cluster sampling, SRS and adaptive SRS sampling (biased).

```
                Adaptive Cluster Sampling

          Number of Observations:          10

            Population Size:        400

                Adaptive Cluster Sampling

                        Statistics

          Mean Var of Mean       Sum Var of Sum

     1.5727273   1.1470888 629.09091  183534.21

                Simple Random Sampling

                        Statistics

          Mean Var of Mean       Sum Var of Sum

           1.2   1.1656667       480  186506.67

        Adaptive Simple Random Sampling (biased)

                        Statistics

          Mean Var of Mean       Sum Var of Sum

     3.1777778   1.0042994 1271.1111   160687.9
```

Figure 8. Output of adaptive spatial cluster sampling.

Table 2. Table of comparison of the estimators of adaptive sampling, SRS and adaptive SRS

Estimator	$\widetilde{\mu}$	$\bar{y}$	$\bar{y}_{AD}$
Mean	1.57	1.20	3.17
Total	629	480	1,271
Variance of the mean	1.147	1.165	1.004
Variance of the total	183,520	186,506	160,687

5.2 Comparison between Different Population Sizes

In this section, interference variation for the total areas (grids) in the estimates will be checked, i.e., the variation in N. Thus we have simulated different sizes of squares and the initial samples were set to have the same sample as the sample (Thompson, 1990). Thus, we obtained the results in Table 3, with the respective values of the estimators of the mean and their estimated variances.

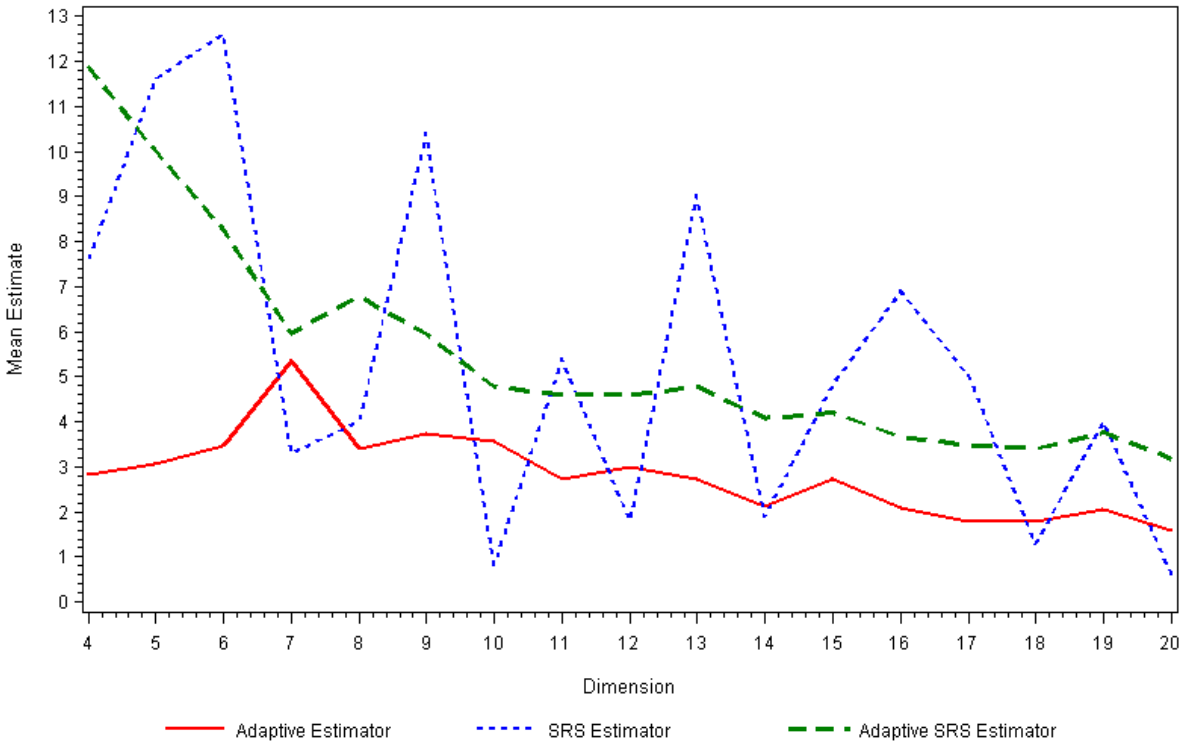

Figure 9. Analysis of the mean when the population size increases (ROOK).

Figure 9 shows how the estimated average is influenced in the case of population variation. We can see that adaptive spatial cluster sampling (solid green line) had the lowest average interference with the change in population, undergoing a decrease with an increase in the size of the regular grid, except when $N = 7$. SRS (blue dotted line) underwent a large change throughout the process. In addition, adaptive SRS sampling - $SRSAD$ (dashed red line) - shows a decrease with a slight increase when $N = 8$.

Table 3. Comparison between the estimators of Mean: adaptive sampling, SRS, adaptive SRS (ROOK)

Matrix	N	$\bar{\mu}$	$\widehat{var}(\bar{\mu})$	$\bar{y}$	$\widehat{var}(\bar{y})$	$\bar{y}_{AD}$	$\widehat{var}(\bar{y}_{AD})$
4x4	16	2.82	1.35	7.60	10.30	11.87	0
5x5	25	3.08	3.43	11.60	42.33	10.00	6.06
6x6	36	3.48	5.99	12.60	62.66	8.26	6.64
7x7	49	5.36	11.57	3.30	3.89	5.96	5.50
8x8	64	3.40	4.70	4.00	7.52	6.78	4.28
9x9	81	3.72	5.41	10.40	47.76	5.93	3.88
10x10	100	3.58	5.13	0.80	0.27	4.76	4.01
11x11	121	2.72	3.01	5.40	13.54	4.61	2.50
12x12	144	2.98	3.85	1.80	1.42	4.61	3.13
13x13	169	2.73	3.10	9.00	36.88	4.76	4.51
14x14	196	2.13	2.31	1.90	1.95	4.08	1.83
15x15	225	2.73	3.22	4.80	10.82	4.20	1.77
16x16	256	2.09	1.91	6.90	22.69	3.67	1.71
17x17	289	1.79	1.44	5.00	15.55	3.49	1.42
18x18	324	1.79	1.45	1.30	0.73	3.40	1.06
19x19	361	2.06	2.03	4.00	7.69	3.76	1.21
20x20	400	1.57	1.15	1.20	1.16	3.18	1.00

Figure 10 represents how the estimate of the variance of the average is influenced in the case of variation in the population. It is apparent that the variance in adaptive sampling is less affected by population size, having a range of 0 to 15, while SRS ranges from 10 to 75 and $SRSAD$ from 0 to 15.

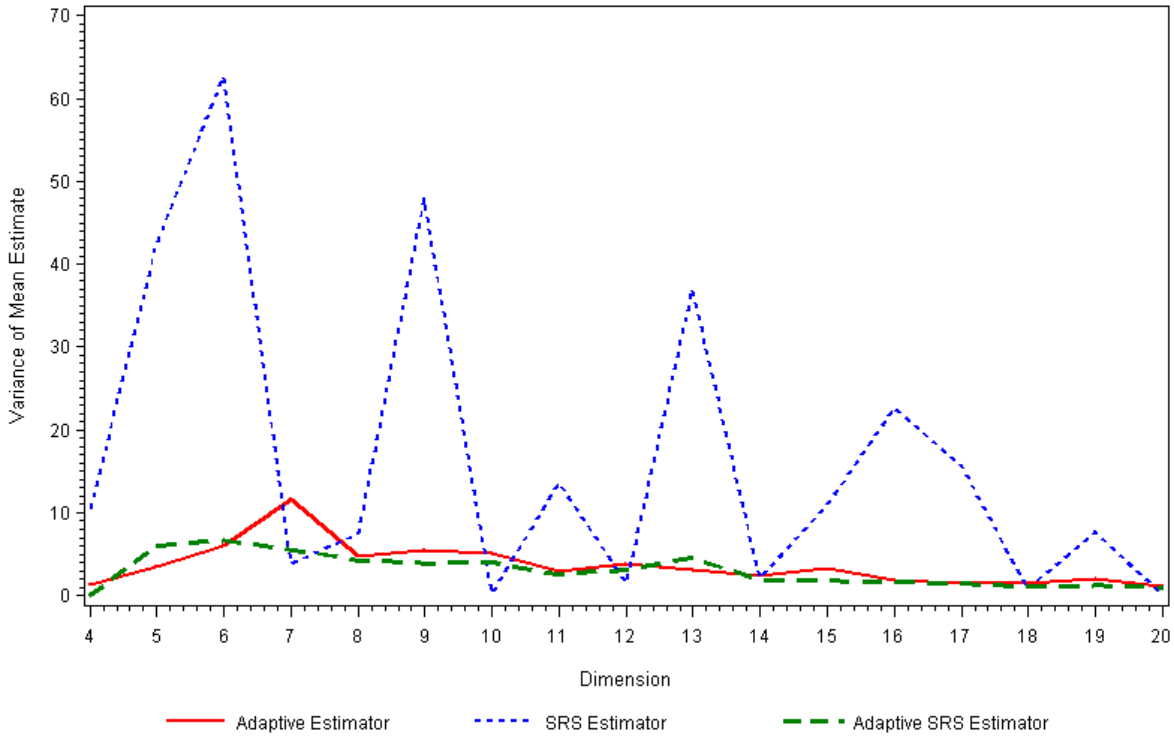

Figure 10. Analysis of the variance when the population size increases (ROOK).

Similarly, we obtain the results in Table 4 for the estimators of the Total. Figure 11 shows the behavior of the total estimator when the population size varies. For this case, the estimator of the total adaptive sampling - $N\mu$ - is

the one with less variation when compared with the other two ($N\bar{y}$ and $N\bar{y}_{AD}$). The variation of the estimator $N\bar{y}$ increases when the population increases.

Table 4. Comparison between the estimators of the total: adaptive sampling, SRS, adaptive SRS (ROOK)

Matrix	N	$N\bar{\mu}$	$N^2\widehat{var}(\bar{\mu})$	$N\bar{y}$	$N^2\widehat{var}(\bar{y})$	$N\bar{y}_{AD}$	$N^2\widehat{var}(\bar{y}_{AD})$
4x4	16	45.20	345.46	121.60	2,637.23	190.00	0
5x5	25	77.02	2,143.20	290.00	26,460.00	250.00	3,786.84
6x6	36	125.28	7,765.19	453.60	81,207.36	297.40	8,605.68
7x7	49	262.97	27,775.32	161.70	9,344.79	291.96	13,224.14
8x8	64	217.60	19,232.25	256.00	30,796.80	434.29	17,532.08
9x9	81	301.72	35,516.02	842.40	313,391.16	480.94	25,485.92
10x10	100	358.33	51,362.50	80.00	2,760.00	476.00	40,096.05
11x11	121	330.16	44,136.80	653.40	198,241.56	558.16	36,596.89
12x12	144	429.60	79,976.56	259.20	29,501.44	664.25	64,971.73
13x13	169	461.13	88,303.16	1,521.00	1,053,343.20	805.57	128,886.39
14x14	196	417.20	89,039.35	372.40	74,896.83	800.80	70,609.67
15x15	225	613.93	162,971.10	1,080.0	548,035.00	946.32	89,533.50
16x16	256	534.75	125,050.12	1,766.40	1,486,863.40	938.67	111,878.68
17x17	289	517.31	120,309.52	1,445.00	1,299,055.00	1,007.97	118,269.46
18x18	324	579.96	151,800.29	421.20	76,980.24	1,103.14	111,112.35
19x19	361	742.75	265,244.93	1,444.00	1,002,424.80	1,358.50	157,809.51
20x20	400	629.09	183,534.21	480.00	186,506.67	1,271.11	160,687.90

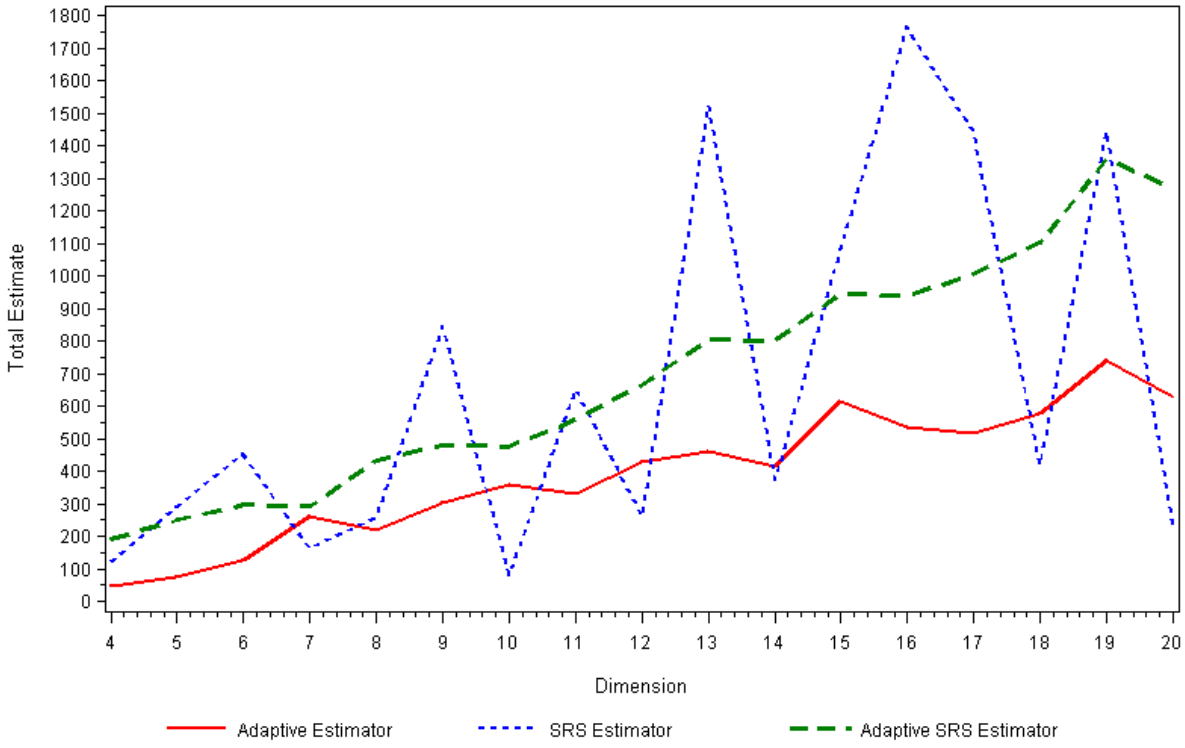

Figure 11. Analysis of the total when the population size increases (ROOK).

Figure 12 represents how the estimator of the variance in the estimator of the total changes with the variation in the population. We can observed that the estimator $N^2\widehat{var}(\bar{\mu})$ and $N^2\widehat{var}(\bar{y}_{AD})$ are closer and $N^2\widehat{var}(\bar{y})$ has large variation throughout the process.

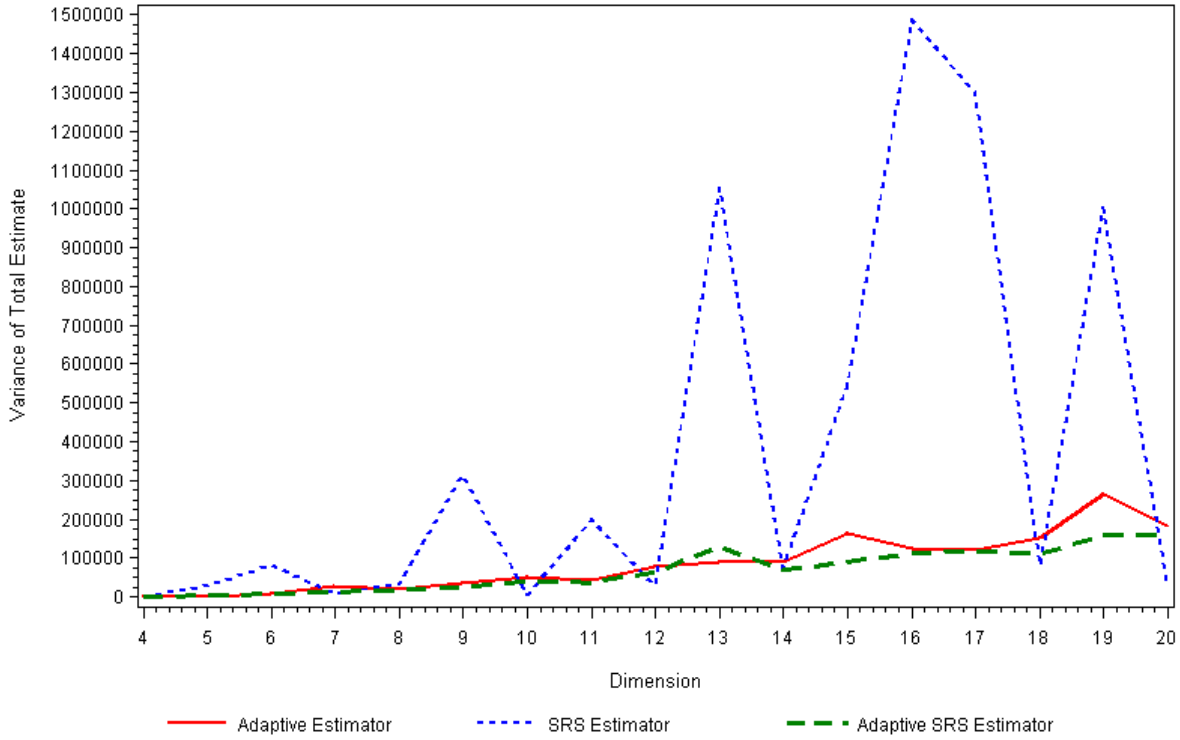

Figure 12. Analysis of the variance of the total when the population size increases (ROOK).

Figure 13 represents how the estimator of the variance in the estimator of the total changes with the variation in the population, with the largest values of $N^2\widehat{var}(\bar{y})$ removed. Notably, the variance is not constant as in Figure 12, showing disorganized growth for SRS and similar growth in the other two cases.

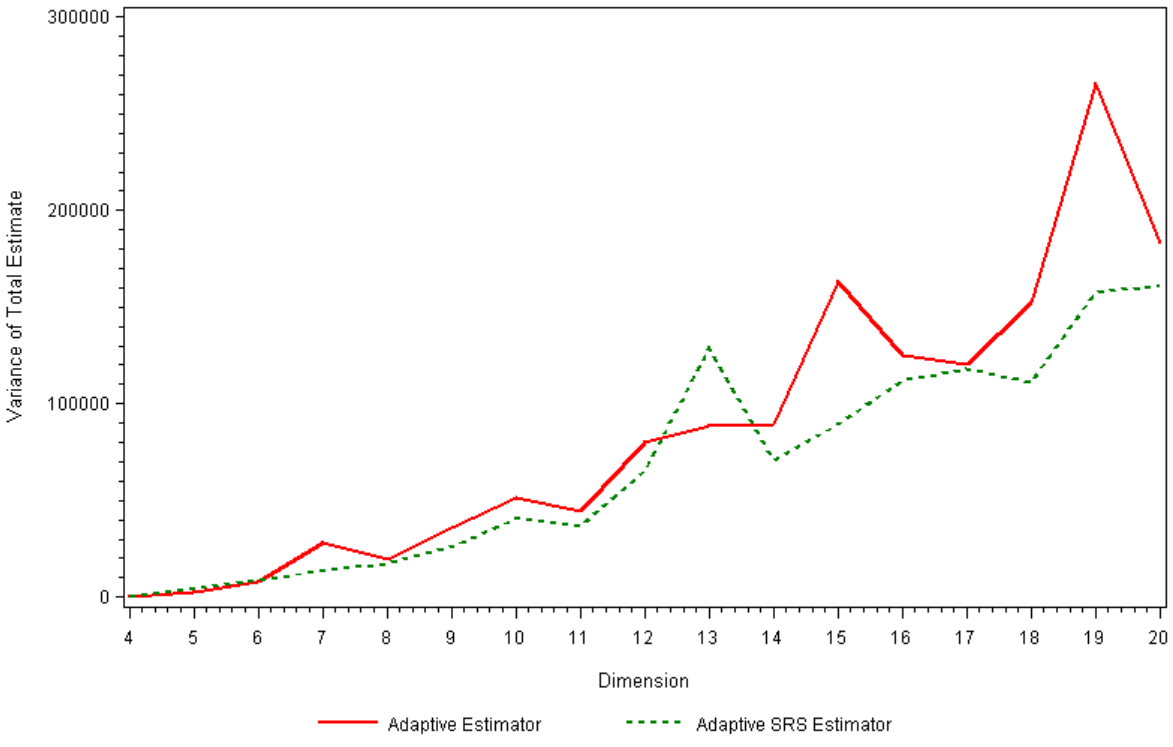

Figure 13. Analysis of the variance of the total when the population size increases (ROOK).

5.3 Stratified Adaptive Spatial Cluster Sampling

Thompson (1990) shows an example of how the stratified adaptive spatial cluster sampling technique works and compares the obtained results when considering whether the boundaries are present between the stratum. Initially, a regular grid is drawn on top of the area to be surveyed, and then, n units (squares) are selected by the SRS method.

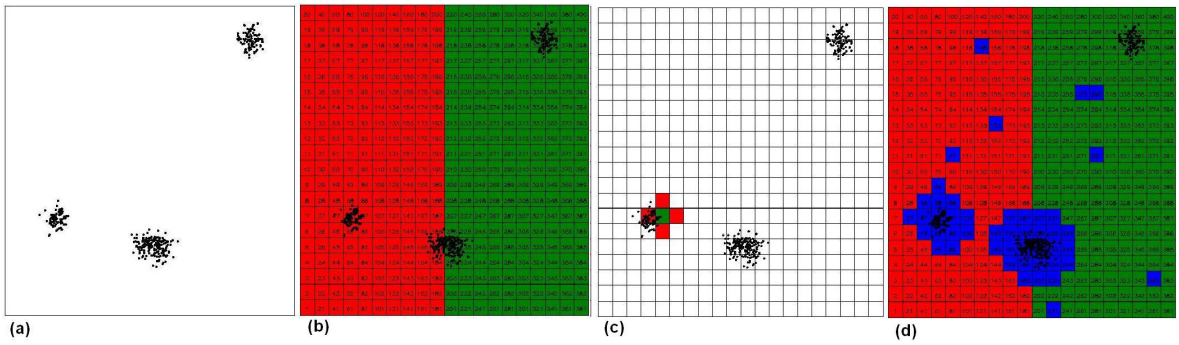

Figure 13. Stratified adaptive spatial cluster sampling.

The number of objects found in the analyzed area in Figure 14 (a) is 397 elements, within a total of $N = 400$ squares. Thus, it follows that the population mean is $\mu = \frac{397}{400} = 0.9925$. For this example, the region was divided into two strata, where for SRS, an initial $n = 10$ elements was selected with stratified adaptive spatial cluster sampling with equal sizes in each stratum. In stratum 1, we see a total of $N = 200$ squares and $n = 5$ elements for the initial sampling units, whereas in stratum 2, the others are $N = 200$ squares and $n = 5$, the total for the area.

As an example of adaptive spatial cluster sampling, the unit satisfies the condition if in each selected square one or more elements of interest is found. Because this condition is verified, it selects its neighbors. The neighborhood of each unit includes all adjacent units. Thus, a neighborhood can be analyzed in two ways: ignoring the existing boundary between strata to select the neighbors of a unit, or considering this limit.

In the first case, a unit to be selected as an element and that is in the square that has contact with the division of the stratum will have four neighbors - top, bottom, right and left - regardless of whether it is in a different stratum. Thus, the value of w'_{hi} for the estimator $\widetilde{\mu}'$, which ignores the limits of the strata, is zero for all units that do not satisfy the condition.

The first network intersect stratum 1, given in Figure 14 (d), and has a value of $w'_{11} = \frac{96}{6} = 16$. For the second network of intersection, the value is given by $w'_{12} = \frac{78}{5} = 15.6$, based only on the units of stratum 1. Thus, there is no intersection in stratum 2. Therefore, the estimate of the mean of the population and the estimated variance of $\widetilde{\mu}'$, given by Equations (35) and (37), respectively, is:

$$\widetilde{\mu}' = \frac{1}{400}\left[\frac{200}{5}(16 + 15.6 + 0 + 0 + 0) + \frac{200}{5}(0 + 0 + 0 + 0 + 0)\right] = 3.16$$

$$\widehat{var}(\widetilde{\mu}') = \frac{1}{400^2}\left[\frac{200(200-5)(74.9)}{5} + 0\right] = 3.65$$

where 74.9 is the variance of the five numbers $(16; 15.6; 0; 0; 0)$.

In the second case, a unit to be selected as an element and that is in the square that has contact with the division of the stratum will have three neighbors: top, bottom, right (or left), depending on whether this is in a different stratum. Thus, to calculate the estimator $\widetilde{\mu}$ (32), it is used for the same stratum $\frac{n_h}{N_h}$.

This obtains the variables w_{hi}, i.e., $w_{11} = \frac{96}{6} = 16$ for the first network and $w_{12} = \frac{192}{11} = 17.45$ for the second. The estimate for the mean and its variance given by Equations (32) and (33), respectively, is:

$$\widetilde{\mu} = \frac{1}{400}\left[\frac{200}{5}(16 + 17.45 + 0 + 0 + 0) + 0\right] = 3.35$$

$$\widehat{var}(\widetilde{\mu}) = \frac{1}{400^2}\left[\frac{200(200-5)(84.2)}{5} + 0\right] = 4.10$$

where 84.2 is the variance given by the five values of w_{1i}.

Table 5 presents the calculated estimates and Figure 15 shows the output of SAS; it appears that there is no significant difference between the estimators of the mean and variance of the estimated average if the boundaries of the strata are considered.

Table 5. Comparison between the estimators of stratified adaptive spatial cluster sampling.

Estimators	No crossing stratum boundaries	Crossing stratum boundaries
Estimators of the mean	3.16	3.35
Variance estimate of the mean	3.65	4.10

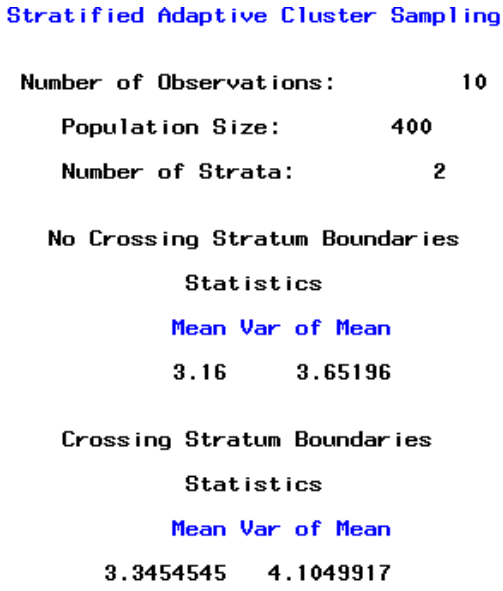

Figure 15. Output of the stratified adaptive spatial cluster sampling.

6. Final Remarks

This study shows that adaptive spatial cluster sampling suffers less variation between SRS and adaptive SRS, which demonstrates that it is a biased estimate. In Section 5, a comparison between different population sizes indicated that adaptive spatial sampling by cluster suffers less variation in the estimators than the other two samples.

Stratified adaptive spatial cluster sampling showed no significant difference between the estimators of the mean, or the estimated average, regardless of whether we consider the limits of the strata variances.

In conclusion, it follows that the computational algorithm for adaptive spatial sampling in this work is important, as this new technique has a variety of applications and users thus far do not have a computational tool to use it.

References

Brown, J. A. (1994). The application of adaptive cluster sampling to ecological studies. *Statistics in ecology and environmental monitoring, 2*, 86 C 97.

Brown, J. A. (1996). The relative effciency of adaptive cluster sampling for ecological surveys. Faculty of Infornation and Mathematical Sciences.

Chang, M. (2008). Adaptive Design Theory and Implementation Using SAS and R. Chapman and Hall/CRC Biostatistics Series.

Chang, M. (2009). Adaptive design theory and implementation using SAS and R. CRC Press.

Cochran, W. G. (1977). Sampling Techniques (3rd ed.). Wiley.

Domingo, C., Gavald'a, R., & Watanabe, O. (2002). Adaptive sampling methods for scaling up knowledge discovery algorithms. *Discovery Science Lecture Notes in Computer Science, 6*(2), 131-152. http://dx.doi.org/10.1023/A:1014091514039.

Jain, A., & Chang, E. Y. (2004). Adaptive sampling for sensor networks. Proceedings of the first workshop on data management for sensor networks, 10-16.

Khan, A., & Muttlak, H. A. (2002). Adjusted two-stage adaptive cluster sampling. *Environmental and Ecological Statistics, 9*, 111-120. http://dx.doi.org/10.1023/A:1013723226430.

Kunigami, G. (2010, novembro). Ponto dentro de polígono. Technical report, Unicamp.

Ramsey, S. K. T. F. L., & Seber, G. A. F. (1992). An adaptive procedure for sampling animal populations. *International Biometric Society, 48*(4), 1195-1199.

Satyanarayana, A., & Davidson, I. (2005). A dynamic adaptive sampling algorithm (dasa) for real world applications: Finger print recognition and face recognition. *Foundations of Intelligent Systems (3488)*, 631-640. http://dx.doi.org/10.1007/11425274_65.

Seber, G. A. F. (1986). A review of estimating animal abundance. *International Biometric Society, 42*(2), 267-292.

Sengupta, R. N., & Sengupta, A. (2011). Some variants of adaptive sampling procedures and their applications. *Computational Statistics and Data Analysis, 55*, 3183-3196. http://dx.doi.org/10.1016/j.csda.2011.05.020.

Stein, A., & Ettema, C. (2003). An overview of spatial sampling procedures and experimental design of spatial studies for ecosystem comparisons. *Agriculture, Ecosystems and Environment, 94*(1), 31-47. http://dx.doi.org/10.1016/S0167-8809(02)00013-0.

Thompson, S. K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association, 85*(412), 1050-1059.

Thompson, S. K. (1991). Stratified adaptive cluster sampling. *Biometrika Trust, 78*(2), 389-397.

Thompson, S. K. (2011). Adaptive sampling. Technical report, Simon Fraser University.

Thompson, S. K., & Seber, G. A. F. (1996). Adaptive sampling. Wiley.

Waldispühl, J., & Ponty, Y. (2011). An unbiased adaptive sampling algorithm for the exploration of rna mutational landscapes under evolutionary pressure. *Research in Computational Molecular Biology Lecture Notes in Computer Science, 6577*, 501-515. http://dx.doi.org/10.1007/978-3-642-20036-6_45.

Yu, H., Jiao, Y., Su, Z., & Reid, K. (2012). Performance comparison of traditional sampling designs and adaptive sampling designs for fishery-independent surveys: A simulation study. *Fisheries Research, 113*(1), 173-181. http://dx.doi.org/10.1016/j.fishres.2011.10.009.