Received: 14 August 2023 | Revised: 9 July 2024 | Accepted: 22 July 2024

DOI: 10.1002/env.2878

RESEARCH ARTICLE

WILEY

How to find the best sampling design: A new measure of spatial balance

Wilmer Prentius | **Anton Grafström**

Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden

Correspondence
Wilmer Prentius, Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-90183 Umeå, Sweden.
Email: wilmer.prentius@slu.se

Abstract

We present a novel measure to assess the spatial balance of a sample by utilizing the balancing equation, which captures the balance between the sample units and their neighbours. Spatially balanced samples are desirable as they may reduce the variance of an estimator of a population parameter. If the auxiliary variables we employ to spread the sample possess high explanatory power for the variable(s) of interest, the resulting reduction in variance can be substantial. An advantageous aspect of using auxiliary variables is that their availability is not contingent upon the sampling effort. Therefore, we can assess and compare sampling designs before committing resources to full-scale surveys. By comparing the proposed measure with commonly used measures of spatial balance, we ascertain that our measure consistently yields meaningful insights regarding the spatial balance of samples. Consequently, our measure can effectively differentiate between various designs when planning a survey, evaluate the potential gains from replacing an existing sample, and determine which non-responding units would contribute the most to enhancing the set of responding units.

KEYWORDS

auxiliary variables, design-based sampling, spatial balance, spatial sampling, statistical measure

1 | INTRODUCTION

The objective of a survey is to gather information about a specific phenomenon within a population, whether finite or infinite, by measuring only a subset, or sample, of that population. The underlying idea is that the information obtained from the sample is sufficiently accurate, while being more cost-effective compared to surveying the entire population.

If a probability sample is used, an unbiased estimator may be constructed. The accuracy of this estimator is often described by its variance, which is determined by the sampling design. The cost of the survey is typically associated with the expected sample size resulting from the sampling design.

For a survey with a fixed budget, that is, with a restriction on the sample size, a sampling design that yields a lower variance of the estimator is preferable to one with a higher variance, assuming both estimators are unbiased. A common way to reducing variance is to use some auxiliary information, available for all units in the population, in the sampling design. For instance, when sampling a population where the gender is known in advance, a stratified sampling design can be used to obtain a sample with the same gender proportions as the population. Similarly, if age is known, an ordered systematic sampling design can be employed to achieve a sample with a similar age distribution as the population.

If gender or age is related to the variables we aim to measure, these sampling designs typically produce samples that better represent the variables of interest, resulting in lower variances compared to designs that ignore this auxiliary information.

In recent years, several sampling designs using auxiliary information have been proposed: General random-tessellation stratified design (Stevens Jr. & Olsen, 2004), the cube method (Deville & Tillé, 2004), spatially correlated Poisson sampling (Grafström, 2012), the local pivotal method (Grafström et al., 2012), balanced acceptance sampling (Robertson et al., 2013), probability function proportional to within sample distance design (Benedetti & Piersimoni, 2017), and weakly associated vectors (Jauslin & Tillé, 2020), among others.

Some of these designs operate under the principle of spreading the sample units in the auxiliary space, generating *well-spread* samples. Similar to the ordered systematic sample design, such designs are generally better at capturing the distribution of the auxiliary variables in the population, when compared to simple random sampling (Grafström & Schelin, 2014). When the sample distribution of the auxiliary variables mirrors the distribution in the population, the sample can be considered spatially balanced.

A spatially balanced, or well-spread, sample is desirable for multiple reasons. Firstly, a well-spread sample is approximately balanced, meaning that the Horvitz-Thompson estimates of the totals of the auxiliary variables are close to the true totals (Grafström & Lundström, 2013). This can have a variance-reducing effect when the auxiliary variables are related to the variable(s) of interest (Grafström & Schelin, 2014). Secondly, if the collected information is to be used in a model, a spatially balanced sample can minimize the risk of large gaps in the collected data, as all parts of the auxiliary space are more likely to be represented in the sample.

If spatially balanced samples are desirable, the question becomes how to determine if a sampling design produces spatially balanced samples. In Section 2, we reiterate the concepts and definitions of spatial balance. A new measure of spatial balance is introduced in Section 3, and is in Section 4 compared to the Voronoi spatial balance measure (Stevens Jr. & Olsen, 2004), and the spatial balance measure based on Moran's I (Tillé et al., 2018). The measures are evaluated by simulation in Section 5, before some final remarks are given in Section 6.

2 | SPATIAL BALANCE

Grafström and Schelin (2014) proposes the following definition of spatial balance: For a finite population U, a sample is well-spread or spatially balanced if, for every coherent subset $U^* \subset U$, we have

$$\sum_{i \in U^*} I_i \approx \sum_{i \in U^*} \pi_i, \tag{1}$$

where I_i is the inclusion indicator and π_i is the inclusion probability of unit i. A subset is considered coherent if it exists within a convex region in auxiliary space. Furthermore, Grafström and Schelin (2014) states that: "A representative sample from a population will be a scaled-down version of the entire population, where all different characteristics of the population are present. With equal inclusion probabilities, a sample well spread in the space spanned by the auxiliary variables corresponds to a representative sample." Using the nomenclature of Kruskal and Mosteller (1979a, 1979b, 1979c), the intended meaning of the term *representative sample* is therefore that of *miniature of the population*.

As the coherency of U^* is defined by U^* being a convex region in auxiliary space, the definition of a representative sample could be rephrased as the weighted population distribution $F_{\mathbf{x}}(U^*)$ being close to the weighted sampling distribution $\widehat{F}_{\mathbf{x}}(U^*)$ for some auxiliary variable $\mathbf{x}$ and every coherent subset U^*. If we denote the set total over $\mathbf{x}$ as $t_{\mathbf{x}}(A) = \sum_{i \in U} I(i \in A) x_i$, then the weighted population and sampling distributions can be described as

$$F_{\mathbf{x}}(U^*) = \frac{t_{\mathbf{x}}(U^*)}{t_{\mathbf{x}}(U)}, \quad \text{and} \quad \widehat{F}_{\mathbf{x}}(U^*) = \frac{t_{\mathbf{x}}(U^* \cap S)}{t_{\mathbf{x}}(S)}, \tag{2}$$

where S denotes the set of units included in the sample. When $\mathbf{x} = \mathbf{1}$, (2) becomes the empirical distribution of the population and sample, $F_{\mathbf{1}}$ and $\widehat{F}_{\mathbf{1}}$ respectively.

As $\widehat{F}$ does not account for the design, we introduce the design-weighted empirical distributions

$$G_{\mathbf{x}}(U^*) = F_{\mathbf{x}}(U^*), \quad \text{and} \quad \widehat{G}_{\mathbf{x}}(U^*) = \frac{1}{t_{\mathbf{x}}(U)} \sum_{i \in U} I(i \in U^*) \frac{x_i}{\pi_i} I_i. \tag{3}$$

The Grafström and Schelin (2014) definition of spatial balance (1) becomes

$$\hat{G}_{\pi}(U^*) \approx G_{\pi}(U^*). \tag{4}$$

If $\pi_i = n/N$, then $G_{\pi} = F_1$ and $\hat{G}_{\pi} = \hat{F}_1$.

While F and $\hat{F}$, or G and $\hat{G}$ might be equal for some coherent subsets U^*, they will differ for most, which is part of the critique given to the term *representative sample* being ambiguous (Kruskal & Mosteller, 1979b). In this paper, we use the term *spatially balanced sample* for samples that, to some degree, satisfy $\hat{G}_{\mathbf{x}}(U^*) \approx G_{\mathbf{x}}(U^*)$ for all available auxiliary variables $\mathbf{x}$.

For a without replacement sampling design $(\mathcal{S}, P)$ on a finite population, the mean spatial balance M is defined through a spatial balance measure B as

$$M(\mathcal{S}, P) = \sum_{S \in \mathcal{S}} P(S) B(S),$$

where $\mathcal{S}$ represents the set of possible samples, and P the probability distribution of the design. In practice, it is often infeasible to calculate $M(\mathcal{S}, P)$, due to the size of $\mathcal{S}$, or because P may be unknown. Hence, the spatial balance is often approximated through Monte Carlo simulation:

$$M^*(\mathcal{S}) = \frac{1}{R} \sum_{r=1}^{R} B(S_r), \quad S_r \in \mathcal{S}, \tag{5}$$

which approaches $M(\mathcal{S}, P)$ as R goes to infinity.

Another important property in sampling is defined through the balancing equation (Tillé, 2006)

$$\sum_{i \in U} \frac{x_i}{\pi_i} I_i = \sum_{j \in U} x_j,$$

or equivalently

$$\hat{G}_{\mathbf{x}}(U) = G_{\mathbf{x}}(U).$$

Sampling designs that satisfy the balancing equation can be very effective, that is, have a relatively low variance compared to other sampling designs, if the variable of interest is a near linear combination of the auxiliary variables (Deville & Tillé, 2005).

3 | A NEW MEASURE OF SPATIAL BALANCE

The prevalent measure of spatial balance is the Voronoi spatial balance measure, introduced by Stevens Jr. and Olsen (2004). Let $\{U_i\}_{i \in S}$ denote the Voronoi partitioning of the population U around the sample units S, such that U_i is the set of population units closer to unit $i \in S$ than any other sample unit. The Voronoi spatial balance measure is then defined by the squared discrepancy of (4) as

$$B_{\mathrm{VO}}(S) = |S| \sum_{i \in S} \left(\hat{G}_{\pi}(U_i) - G_{\pi}(U_i) \right)^2,$$

where $|S|$ denotes the size of the set S. A sampling design that produces fixed sized samples is balanced on the inclusion probabilities $\boldsymbol{\pi}$, that is, satisfies the balancing equation

$$\hat{G}_{\pi}(U) = G_{\pi}(U).$$

Thus, the Voronoi spatial balance measure can be seen as the ability of the neighbourhood, through the sum of their inclusion probabilities, to represent the number of selected units within it.

An alternative definition of the spatial balance (1) could be formed by considering the balancing equation for an auxiliary variable $\mathbf{x} = \mathbf{1}$, that is, balancing on the population size. A fixed sized sample would then be considered spatially balanced if, for every coherent subset $U^* \subset U$, the selected units represents the assigned number of population units; the ability of a selected unit, through its inclusion probability, to represent the size of its neighbourhood.

Other than measuring if the size of the neighbourhoods matches the sample units (or vice versa), the spatial balance measure could account for the local balance of the auxiliary variables, that is, ability of the selected unit to represent its neighbourhood in the auxiliary space. Let

$$d_{\mathbf{x}}(U^*) = t_{\mathbf{x}}(U)\left[\hat{G}_{\mathbf{x}}(U^*) - G_{\mathbf{x}}(U^*)\right]$$

denote a measure of the discrepancy of the balancing equation for some auxiliary variable $\mathbf{x}$ and neighbourhood U^*, and let $\mathbf{d}_{\mathbf{X}}(U^*)$ denote the p-length vector of the measures of discrepancies amongst all p centered auxiliary variables in $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_p]$. The proposed measure of spatial balance, the local spatial balance measure, can then be defined as

$$B_{LB}(S) = \sqrt{\frac{1}{|U|}\sum_{i \in S} \mathbf{d}_{\mathbf{X}}(U_i)^\intercal \mathbf{Q}^{-1} \mathbf{d}_{\mathbf{X}}(U_i)}, \tag{6}$$

where $\mathbf{Q} = \mathbf{X}^\intercal\mathbf{X}$ is the Gram matrix of $\mathbf{X}$. The measure, related to the Mahalanobis distance, measures the squared discrepancies in the neighbourhoods around the sample units, with respect to $\mathbf{Q}$. The use of $\mathbf{Q}$ instead of the covariance matrix is proposed to allow the inclusion of balancing with respect to the size of the neighbourhood, that is, $\mathbf{X} = [\mathbf{1}, \mathbf{x}_1, \dots, \mathbf{x}_p]$.

4 | COMPARISON BETWEEN CURRENT SPATIAL BALANCE MEASURES

A drawback of the Voronoi spatial balance measure, as well as the proposed balance measure, is that it does not operate on a fixed scale. In order to interpret whether a design is more spatially balanced than another, we need to compare the sampling distributions of each design, which will be infeasible to do for anything other than very small populations. Another alternative is to compare the mean spatial balance of the designs, estimated through Monte Carlo simulation as in (5). Often, when evaluating if a design is efficient, the mean spatial balance of the design is compared to a baseline design, such as the Simple Random Sampling (SRS) design.

The spatial balance measure based on Moran's I, introduced by Tillé et al. (2018), tries to overcome this problem by fixing the measure to a range $[-1, 1]$, where -1 implies perfect spatial balance, and 1 maximum concentration. The measure is described as a weighted correlation between the inclusion indicators and the average value of the inclusion indicators of its neighbourhood (Tillé et al., 2018). The weights are decided by the inclusion probabilities (Jauslin & Tillé, 2020).

Consider a population as in Figure 1, where we have three similar clusters, each with three units. The clusters are separated so that the distance between units within a cluster is less than the distance between any two units of different clusters.

A design that selects well-spread samples is the stratified design, drawing a unit from each cluster independently with equal probabilities. The upper part of Table 1 shows the possible samples of such a design, together with the spatial balance measures for each of these samples. The mean spatial balance measures for the design are 0.0, −1.0, and 0.41 for the Voronoi spatial balance measure (VO), the spatial balance measure based on Moran's I (MI), and the proposed spatial balance measure (LB) respectively. Neither of the VO or MI manages to discriminate between the units within a cluster.

For an SRS (without replacement), that is, a design which allows all possible combinations of samples, we need to consider all the outcomes in Table 1. The mean spatial balance measures for the SRS design are 0.23 (VO), −0.29 (MI), and 0.54 (LB), hence all measures accurately confirms that the stratified design is more well-spread than the SRS design. The MI accurately captures the most clustered sample, but does not distinguish between any other samples. Furthermore, the VO gives somewhat erratic results, when selecting two units from one cluster, and one from another.

For the two one-dimensional populations in Figure 2, the results are given in Table 2 for SRS samples drawn with a sample size of 2. Here, it is clear that the VO only considers the sizes of the Voronoi partitioning, not distinguishing between samples with the same partition sizes. Furthermore, while there seems to be an agreement between the measures in Table 2 for the equispaced population (a), the effect of the Voronoi partitioning of VO and LB is clearer for population (b). For instance, a strong disagreement occurs for the sample $a_i c_j$ (population b). VO and LB considers this as one of the

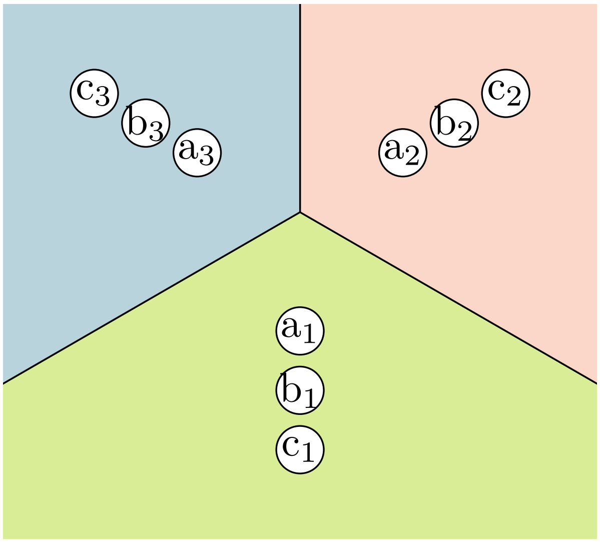

FIGURE 1 A population of three clusters 1, 2, 3, each with three units *a*, *b*, *c*. The distance between units within a cluster is smaller than the distance between any units of different clusters.

TABLE 1 The spatial balance measures for samples of size 3 from the population in Figure 1.

Sample	#	VO	MI	LB
(a) Stratified samples				
aaa	1	0	−1	0.26
aab	3	0	−1	0.21
aac	3	0	−1	0.26
abb	3	0	−1	0.15
abc	6	0	−1	0.21
acc	3	0	−1	0.26
bbb	1	0	−1	0.00
bbc	3	0	−1	0.15
bcc	3	0	−1	0.21
ccc	1	0	−1	0.26
Mean	27	0	−1	0.20
(b) Clustered samples				
$a_i b_i c_i$	3	0.89	1	0.99
$a_i b_i a_j$	6	0.52	0	0.62
$a_i b_i b_j$	6	0.07	0	0.61
$a_i b_i c_j$	6	0.07	0	0.63
$a_i c_i a_j$	6	0.50	0	0.67
$a_i c_i b_j$	6	0.17	0	0.67
$a_i c_i c_j$	6	0.17	0	0.68
$b_i c_i a_j$	6	0.52	0	0.73
$b_i c_i b_j$	6	0.52	0	0.76
$b_i c_i c_j$	6	0.30	0	0.79
Mean	84	0.23	−0.29	0.54

Note: The samples have been divided into well-spread (a) and clustered (b) samples, where the stratified samples select one unit from each group, and clustered samples consists of any other combination of units from groups $i \neq j$. The mean spatial balance measures given in (b) also considers the possible samples in (a).

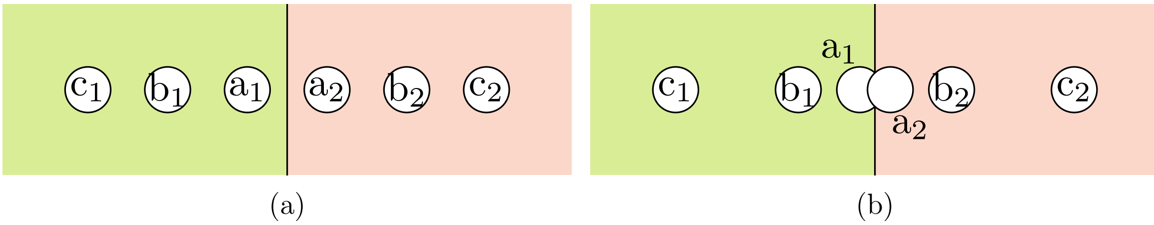

FIGURE 2 Two populations of six units each on a one dimensional line, where (a) consists of equispaced units, and the units of (b) have increasing distances towards the edges.

TABLE 2 The spatial balance measures for samples of size 2 from the populations in Figure 2, units from groups $i \neq j$.

Sample	#	VO	MI	LB
(a)				
a_ia_j	1	0.00	—	0.41
a_ib_j	2	0.11	−0.86	0.34
a_ic_j	2	0.11	−1.00	0.48
b_ib_j	1	0.00	−1.00	0.00
b_ic_j	2	0.03	−0.71	0.34
c_ic_j	1	0.00	−0.50	0.41
a_ib_i	2	0.11	0.17	0.59
a_ic_i	2	0.11	−0.16	0.68
b_ic_i	2	0.44	0.71	0.96
Mean	15	0.12	—	0.51
(b)				
a_ia_j	1	0.00	−0.50	0.47
a_ib_j	2	0.11	−0.61	0.39
a_ic_j	2	0.44	−1.00	0.74
b_ib_j	1	0.00	−1.00	0.12
b_ic_j	2	0.11	−0.50	0.50
c_ic_j	1	0.00	—	0.58
a_ib_i	2	0.11	0.00	0.50
a_ic_i	2	0.44	−0.25	0.79
b_ic_i	2	0.44	0.25	0.92
	15	0.22	—	0.59

Note: The MI is undefined for two samples.

worst spatially balanced samples, due to c_j being the sole unit in one Voronoi partition of the two partitions, whereas MI considers it as a perfectly spatially balanced sample, as no other unit in the immediate neighbourhood of the sample units have been selected to the sample.

Notably, the MI becomes undefined in two circumstances: For the population in Figure 2a, the measure is undefined when evaluating a sample to which each unit in the population assigns equal weight, or in the case of an equal probability design, when each unit in the population considers exactly n units in the sample to be a neighbour. For the population given in Figure 2b, the measure is undefined when evaluating a sample which contains no units that any other unit in the population considers a neighbour.

A practical issue arises when a population unit has the same distance to multiple units. Both the MI and the VO handles this by sharing the unit (or weights) equally between the ties, and we suggest the LB to function similarly.

5 | SIMULATION

In order to evaluate the proposed measure of spatial balance (LB) against the Voronoi measure (VO) and the measure based on Moran's I (MI), and to see how the measures perform for different populations and sample sizes, a simulation experiment was conducted. Three artificial populations were constructed, seen in Figure 3, with approximately 1000 units each. The first population was constructed through a Poisson cluster process with 100 parent positions spawning a mean of 10 units each, resulting in a population of $N = 975$; the second from a uniform distribution ($N = 1000$); the third through a hexagon pattern ($N = 957$).

From each population, 100,000 samples were drawn from each combination of three different sample sizes (50, 100, 200) and two different designs, SRS and the Local Pivotal Method 2 (LPM), where the latter had access to the coordinates as auxiliary variables (Grafström et al., 2012). For each sample, we calculated the spatial balance measures of VO, MI and LB. Implementations for the VO and LB are available through the R package `BalancedSampling` (Grafström et al., 2024), and the MI is available through the R package `WaveSampling` (Jauslin & Tillé, 2022).

In Table 3, the empirical means of the spatial balance measures are presented for the Poisson, uniform and hexagon populations respectively. For the VO and LB spatial balance measures, it is possible to compare two sampling designs by examining the ratio between their respective spatial balance measures (Stevens Jr. & Olsen, 2004). While the LB results are consistent with VO and MI in showing that LPM produces more spatially balanced samples compared to SRS, the

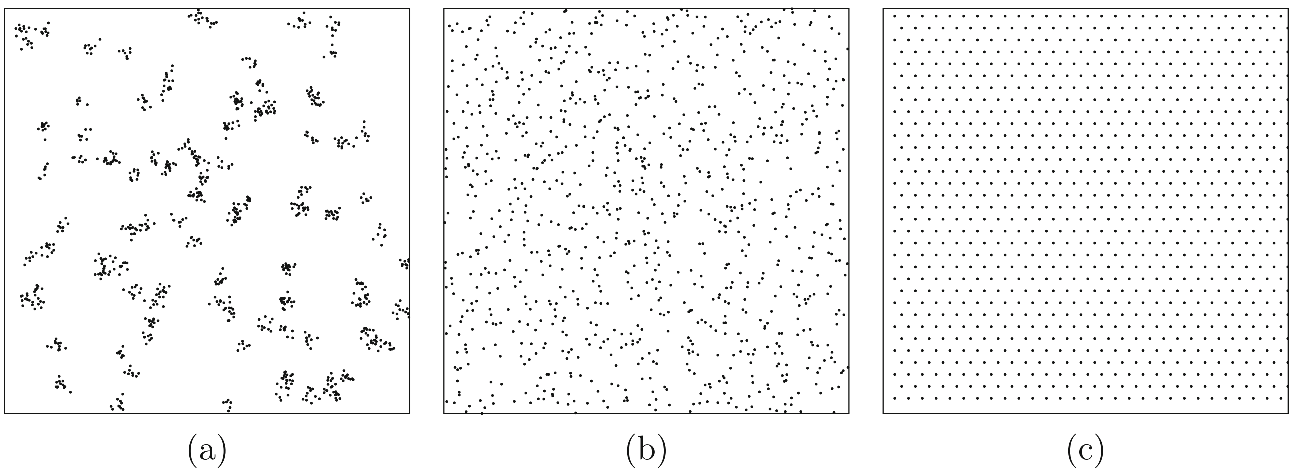

FIGURE 3 Artificial populations used in the simulation: (a) Poisson cluster process, (b) Uniform distribution, (c) Hexagon pattern.

TABLE 3 Mean spatial balance measures for 100,000 samples taken with LPM and SRS with different sample sizes n from the three different populations.

		Poisson			Uniform			Hexagon		
n	**Design**	**VO**	**MI**	**LB**	**VO**	**MI**	**LB**	**VO**	**MI**	**LB**
50	LPM	0.133	−0.285	0.090	0.081	−0.233	0.073	0.070	−0.234	0.070
	SRS	0.482	−0.008	0.169	0.323	−0.007	0.142	0.303	−0.007	0.139
	Ratio	0.275		0.533	0.250		0.513	0.231		0.504
100	LPM	0.131	−0.395	0.061	0.099	−0.313	0.055	0.070	−0.314	0.048
	SRS	0.541	−0.005	0.127	0.331	−0.005	0.101	0.284	−0.005	0.095
	Ratio	0.242		0.485	0.298		0.548	0.246		0.507
200	LPM	0.196	−0.415	0.054	0.126	−0.408	0.043	0.079	−0.421	0.036
	SRS	0.543	−0.003	0.090	0.347	−0.003	0.073	0.261	−0.003	0.064
	Ratio	0.361		0.604	0.362		0.596	0.304		0.556

Note: Ratios of the spatial balance measures of LPM by SRS are calculated for the LB and VO measures.

ratios indicate that VO expresses a higher relative increase in the spatial balance of LPM, whereas the increase for LB is less pronounced.

Furthermore, a smaller set of 1000 samples were taken for each population and design, for increasing sample sizes by 10. The mean spatial balance measures for this simulation is shown in Figure 4, with relative mean spatial balance measures of LPM versus SRS shown for LB and VO in Figure 5. While convergence of this smaller simulation is less likely, it still establishes the general trends of the measures for different samples sizes.

For the first, smaller sample sizes, both the MI and the LB shows a rapid increase of spatial balance, where after the effect of spatial balancing subsides. If a spatially balanced sampling method, such as LPM, is seen as a method that reduces the risk of obtaining worst case samples, this behaviour is what we expect. LPM is very efficient in the beginning, but as sample sizes increases, the risk of choosing a very clustered sample becomes smaller for even the SRS. On the

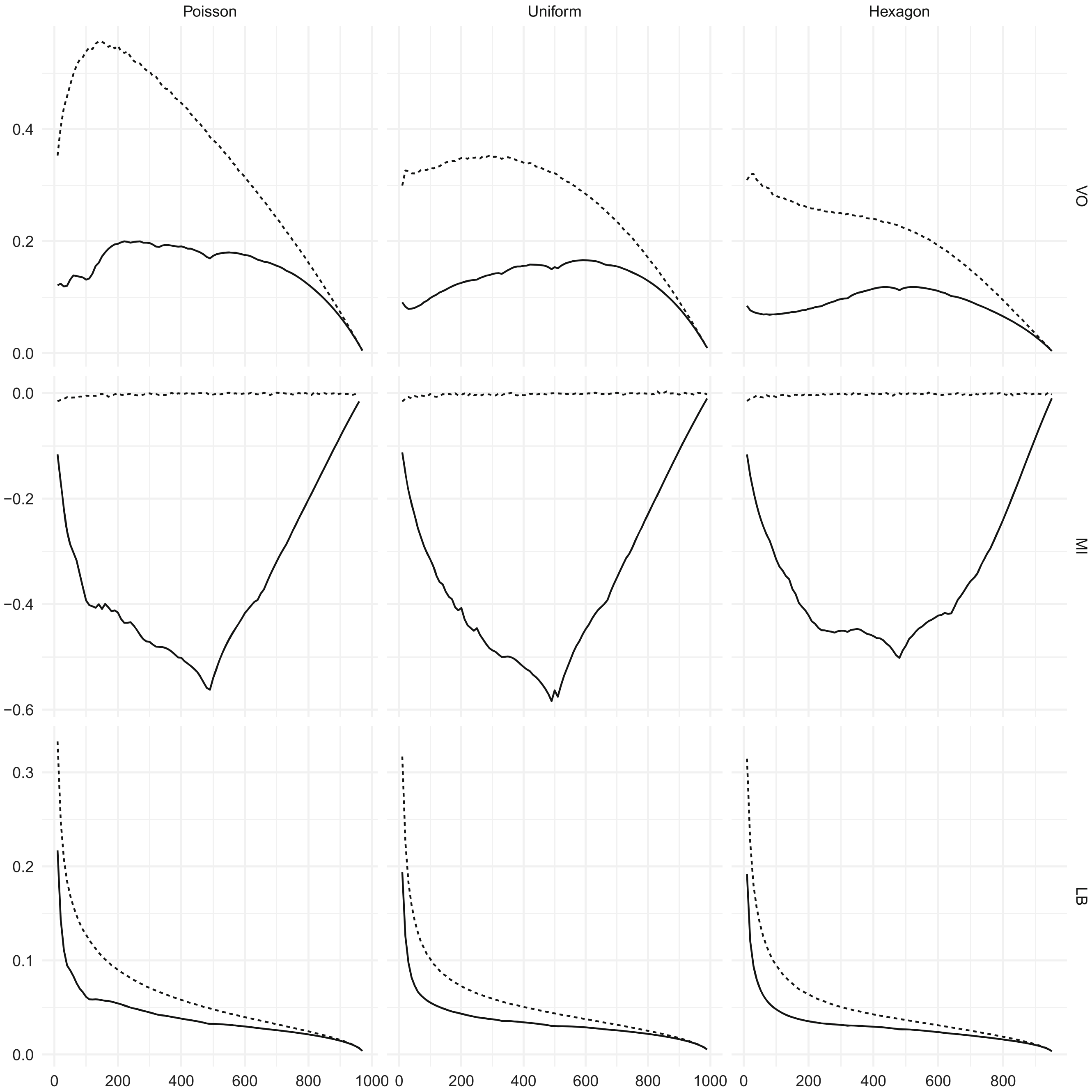

FIGURE 4 Mean spatial balance measures for LPM (solid lines) and SRS (dotted lines), when taking 1000 samples of sample sizes 10, 20, … from the Poisson, uniform, and hexagon populations.

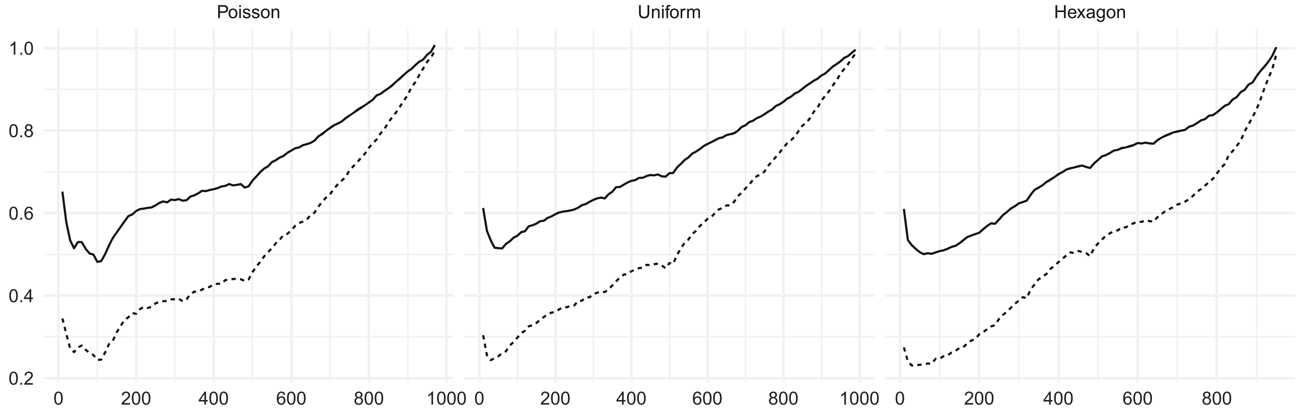

FIGURE 5 Relative mean spatial balance measures for LPM relative to SRS for LB (solid lines) and VO (dotted lines), when taking 1 000 samples of sample sizes 10, 20, … from the Poisson, uniform, and hexagon populations.

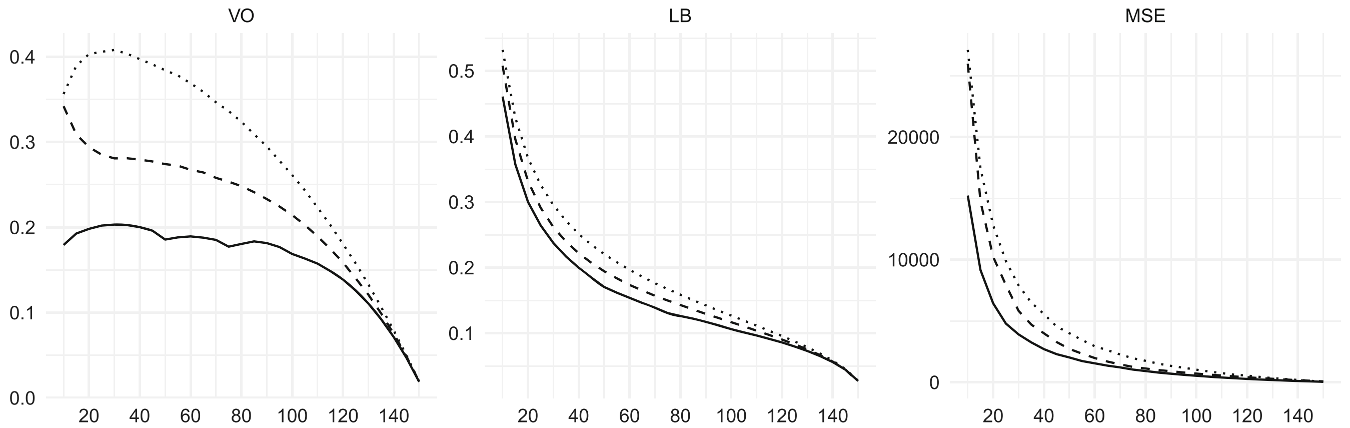

FIGURE 6 Mean spatial balance measures for VO and LB, and the MSEs for the cadmium variable of interest, for the Meuse data and different sample sizes ($n = 10, 15, \ldots$), repeated 10,000 times. Samples drawn by LPM (solid lines), LCUBE (dashed lines), and SRS (dotted lines).

other hand, the VO behaves unintuitively for the Poisson population, as it exhibits a decrease in spatial balance with increasing samples sizes, up until a sample size of about 150. Nevertheless, the ratios in Figure 5 shows a consistency of interpretation between the VO and the LB.

Furthermore, Figure 4 highlights a difference in interpretation. The VO and LB assumes that, as sample size increases towards N, the sample should become more spatially balanced independent of sampling method. The MI, on the other hand, shows that as the sample size increases, from about $N/2$ an onwards, the sample becomes less spatially balanced, as it assumes that a sample of size N is neither spatially balanced or clustered. Together, Figures 4 and 5 indicate that the LB might be a better choice if comparing two samples of different sizes.

In addition to the simulated populations, the methods were evaluated on the Meuse data set from the R package `sp` (Bivand et al., 2013). The data set contains 155 locations and topsoil heavy metal concentrations along the river Meuse. Out of these 155, 2 observations were excluded with missing data.

From the data set, 10,000 samples were taken using LPM, SRS and the Local Cube method (LCUBE) (Grafström & Tillé, 2013). LPM and LCUBE used coordinates, copper concentrations, elevation, and organic matter as auxiliary variables for spreading the sample. Furthermore, LCUBE used copper concentration, elevation and organic matter to balance as balancing variables. The total cadmium levels, the variable of interest, was estimated for each sample, and the estimators MSE reported. The mean spatial balance measures and the MSEs are shown in Figure 6, and the relative measures and MSEs are shown in Figure 7.

Generally, the results in Figure 6 agree with the simulations on the artificial populations, and suggest that the results are consistent in higher dimensions. Furthermore, the benefits of spatially balanced sampling are highlighted in the

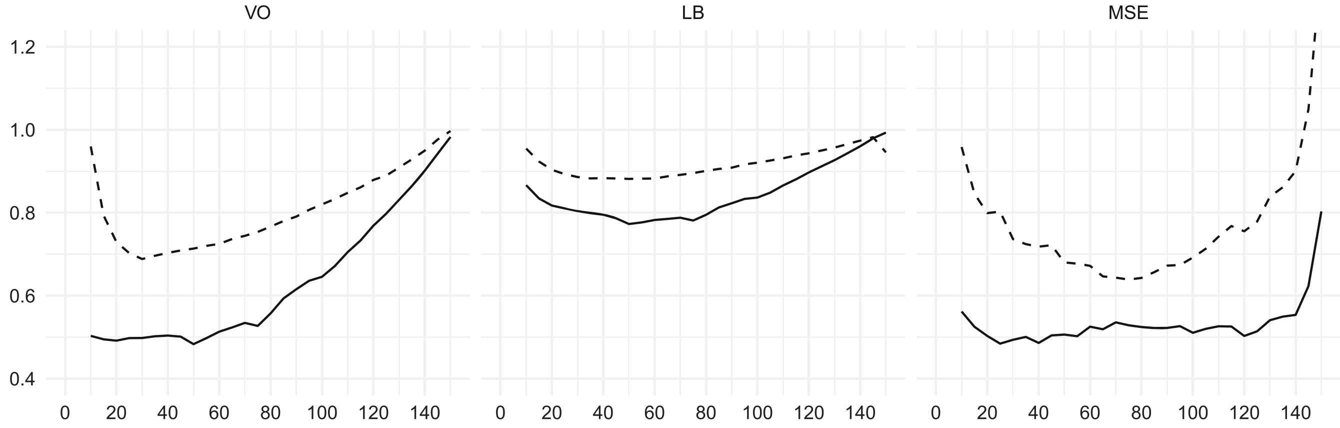

FIGURE 7 Relative mean spatial balance measures for VO and LB, and relative MSEs for the cadmium variable of interest, for the Meuse data and different sample sizes ($n = 10, 15, \ldots$), repeated 10,000 times. LPM relative to SRS (solid lines), and LCUBE compared to SRS (dashed lines).

reduction of MSE for both LPM and LCUBE shown in Figure 7. Notably, the balancing in LCUBE takes away from the ability to spatially balance the samples.

6 | FINAL REMARKS

In this paper, we have presented a new measure of spatial balance. Through examples and simulations, we have shown that the measure has some desirable properties: the measure better discriminates between samples with similar Voronoi partitions, as it measures local balance within the Voronoi polytopes, while also being consistent with respect to having a decreasing form as the sample size increases. The measure could therefore be used to compare the spatial balance of two different samples, taken from the same population. One potential application could be to determine where additional sampling efforts should be taken when following up on non-responses, for example, by sequentially adding the non-respondents who increases the spatial balance the most. This can be compared with the measure of lack of balance in Särndal (2011), which considers only the balance between the respondents and the selected sample, but only requires that the values of the auxiliary variables are known for the sample.

Another potential application is the monitoring of longitudinal surveys. A first sample may have been spatially balanced, but as the auxiliary variables drift over time, it might be necessary to update the sample after time has passed. By monitoring the development of the spatial balance with respect to newly acquired auxiliary data, it is thus possible to gain information of if and when a new sampling effort is needed.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Wilmer Prentius https://orcid.org/0000-0002-3561-290X
Anton Grafström https://orcid.org/0000-0002-4345-4024

REFERENCES

Benedetti, R., & Piersimoni, F. (2017). A spatially balanced design with probability function proportional to the within sample distance. *Biometrical Journal*, *59*(5), 1067–1084. https://doi.org/10.1002/bimj.201600194

Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (Second ed.). Springer. https://asdar-book.org/

Deville, J. C., & Tillé, Y. (2004). Efficient balanced sampling: The cube method. *Biometrika*, *91*(4), 893–912. https://doi.org/10.1093/biomet/91.4.893

Deville, J. C., & Tillé, Y. (2005). Variance approximation under balanced sampling. *Journal of Statistical Planning and Inference*, *128*(2), 569–591. https://doi.org/10.1016/j.jspi.2003.11.011

Grafström, A. (2012). Spatially correlated Poisson sampling. *Journal of Statistical Planning and Inference*, *142*(1), 139–147. https://doi.org/10.1016/j.jspi.2011.07.003

Grafström, A., Lisic, J., & Prentius, W. (2024). BalancedSampling: Balanced and Spatially Balanced Sampling. R package version 2.0.6. https://CRAN.R-project.org/package=BalancedSampling

Grafström, A., & Lundström, N. L. (2013). Why well spread probability samples are balanced. *Open Journal of Statistics*, *3*(1), 36–41. https://doi.org/10.4236/ojs.2013.31005

Grafström, A., Lundström, N. L., & Schelin, L. (2012). Spatially balanced sampling through the pivotal method. *Biometrics*, *68*(2), 514–520. https://doi.org/10.1111/j.1541-0420.2011.01699.x

Grafström, A., & Schelin, L. (2014). How to select representative samples. *Scandinavian Journal of Statistics*, *41*(2), 277–290. https://doi.org/10.1111/sjos.12016

Grafström, A., & Tillé, Y. (2013). Doubly balanced spatial sampling with spreading and restitution of auxiliary totals. *Environmetrics*, *24*(2), 120–131. https://doi.org/10.1002/env.2194

Jauslin, R., & Tillé, Y. (2020). Spatial spread sampling using weakly associated vectors. *Journal of Agricultural, Biological and Environmental Statistics*, *25*(3), 431–451. https://doi.org/10.1007/s13253-020-00407-1

Jauslin, R., & Tillé, Y. (2022). WaveSampling: Weakly Associated Vectors (WAVE) Sampling. R package version 0.1.3. https://CRAN.R-project.org/package=WaveSampling

Kruskal, W., & Mosteller, F. (1979a). Representative sampling, I: Non-scientific literature. *International Statistical Review/Revue Internationale de Statistique*, *47*(1), 13–24. https://doi.org/10.2307/1403202

Kruskal, W., & Mosteller, F. (1979b). Representative sampling, II: Scientific literature, excluding statistics. *International Statistical Review/Revue Internationale de Statistique*, *47*(2), 111–127. https://doi.org/10.2307/1402564

Kruskal, W., & Mosteller, F. (1979c). Representative sampling, III: The current statistical literature. *International Statistical Review/Revue Internationale de Statistique*, *47*(3), 245–265. https://doi.org/10.2307/1402647

Robertson, B. L., Brown, J. A., McDonald, T., & Jaksons, P. (2013). BAS: Balanced acceptance sampling of natural resources. *Biometrics*, *69*(3), 776–784. https://doi.org/10.1111/biom.12059

Särndal, C. E. (2011). The 2010 Morris Hansen lecture. Dealing with survey nonresponse in data collection, in estimation. *Journal of Official Statistics*, *27*(1), 1.

Stevens, D. L., Jr., & Olsen, A. R. (2004). Spatially balanced sampling of natural resources. *Journal of the American Statistical Association*, *99*(465), 262–278. https://doi.org/10.1198/016214504000000250

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

Tillé, Y., Dickson, M. M., Espa, G., & Giuliani, D. (2018). Measuring the spatial balance of a sample: A new measure based on Moran's I index. *Spatial Statistics*, *23*, 182–192. https://doi.org/10.1016/j.spasta.2018.02.001

How to cite this article: Prentius, W., & Grafström, A. (2024). How to find the best sampling design: A new measure of spatial balance. *Environmetrics*, *35*(7), e2878. https://doi.org/10.1002/env.2878