Spatially Balanced Sampling with Local Ranking

B. L. ROBERTSON, O. OZTURK, O. KRAVCHUK, and J. A. BROWN

A spatial sampling design determines where sample locations are placed in a study area so that population parameters can be estimated with good precision. Spatially balanced designs draw samples with good spatial spread and provide precise results for commonly used estimators when surveying natural resources. In this article, we propose a new sampling strategy that incorporates ranking information from nearby locations into a spatially balanced sample. If the population exhibits spatial trends, our simple local ranking strategy can improve the precision of commonly used estimators. Numerical results on several test populations with different spatial structures show that local ranking can improve the performance of a spatially balanced design. To show that local ranking is simple and effective in practice, we provide an example application for the health and productivity assessment of a Shiraz vineyard in South Australia.

Supplementary materials accompanying this paper appear online.

Key Words: Environmental sampling; Quasi-random sampling; Ranked set sampling; Spatial balance.

1. INTRODUCTION

The main objective of a spatial sampling design is to determine where sample locations are placed in a study region so that population parameters can be estimated with good precision. A common goal is to estimate the mean or total, but other population characteristics may also be of interest. Natural resources are spatially distributed populations and their spatial structure should be taken into account when sampling them. The exact spatial structure is usually not known, but a reasonable assumption is that nearby locations tend to have more similar response values than distant ones (Stevens and Olsen 2004). Hence, a simple strategy to improve the precision of commonly used estimators is to spread the sample well over the resource, called spatially balanced sampling (Stevens and Olsen 2004). In this article, we consider drawing spatially balanced (SB) locations from a finite set $\Omega \subset \mathbb{R}^2$.

B. L. Robertson (✉) · J. A. Brown, School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand
(E-mail: *blair.robertson@canterbury.ac.nz*).
O. Ozturk, Department of Statistics, The Ohio State University, Columbus, OH, USA.
O. Kravchuk, School of Agriculture, Food and Wine, University of Adelaide, Adelaide, Australia.

SB designs are known to perform well when sampling natural resources (c.f. Stevens and Olsen 2004; Grafström and Lundström 2013; Grafström and Schelin 2014) and many designs have been proposed. One of the first and most widely used SB design is generalized random tessellation stratified (GRTS, Stevens and Olsen 2004). Stevens and Olsen (2004) define the spatial balance of a sample $\{x_1, x_2, ..., x_n\} \subset \Omega$ using a Voronoi tessellation. The sample is considered spatially balanced if

$$v_i = \sum_{x_j \in \omega_i} \pi_j \approx 1 \quad \text{for all } i = 1, 2, ..., n,$$

where $0 < \pi_i < 1$ denotes the inclusion probability of x_i and ω_i is the set of points from Ω that fall within the Voronoi polygon for x_i, given by

$$\omega_i = \left\{ x \in \Omega : \|x - x_i\| \le \|x - x_j\| \text{ for all } j = 1, 2, ..., n \right\}.$$

GRTS is particularly useful for sampling natural resources because SB over-samples can be drawn if nonresponse is anticipated (Stevens and Olsen 2004; Larsen et al. 2008). Chauvet and Le Gleut (2020) improved the statistical properties of GRTS, and a computationally efficient implementation of GRTS was proposed.

A variety of SB designs have been proposed to draw sample locations from natural resources (c.f. Benedetti et al. 2015; Kermorvant et al. 2019). Several methods map Ω to the real line and then draw sample locations from the one-dimensional frame using, for example, systematic sampling (Stevens and Olsen 1999, 2003, 2004, Theobald et al. 2007, Dickson and Tillé 2016). Benedetti and Piersimoni (2017) proposed a flexible class of designs that draw SB samples based on a within-sample distance or similarity measure. Another group of methods apply a local repulsion strategy to ensure well-spread samples are drawn (Grafström 2011; Grafström et al. 2012; Grafström and Tillé 2013; Grafström and Matei 2018). SB designs based on the Halton sequence (Halton 1960) have been proposed (Robertson et al. 2013, 2017; van Dam-Bates et al. 2018; Robertson et al. 2018, 2021). The Halton sequence is quasi-random and generates evenly spread points with similar spatial properties to a regular lattice; however, unlike a regular lattice, points can be added incrementally with no clumping of points.

In this article, we aim to improve the performance of a SB design by incorporating local ranking information into its sample selection strategy. We consider four SB designs, but the method can be easily incorporated into any design. The first method is GRTS with reverse hierarchical ordering. This ordering ensures that all contiguous GRTS subsamples are SB samples (Stevens and Olsen 2004). The second method is the local pivotal method (LPM, Grafström et al. 2012). LPM is a very flexible design that uses a local repulsion strategy to draw SB samples. LPM can draw unequal probability samples and can spread samples over several auxiliary variables. The third method is the product within distance design (PWD, Benedetti and Piersimoni 2017), which uses the within sample distance as the summary index of the spatial distribution of a random selection criterion. The final method is Halton iterative partitioning (HIP, Robertson et al. 2018), which draws SB samples from a nested partition related to structural properties of the Halton sequence. HIP samples have the property that all contiguous subsamples are SB samples (Robertson et al. 2018).

Ranked set sampling (RSS) is a design that incorporates ranking information into its sample selection strategy. The ranking information creates artificial strata (judgment classes) so that homogeneous units can be grouped together without measuring the response variable directly. Measuring the response variable at points drawn from these artificial strata aims to achieve good coverage over the full range of the response variable. The method was first proposed by McIntyre (1952) and later named RSS by Halls and Dale (1966). The balanced RSS procedure begins by randomly drawing k sets of points from Ω, with each set having k points. The points in each set are then judgment ranked using an efficient and inexpensive method that does not require full measurements of the response variable. Examples of judgment ranking are to use a quick estimate of the response variable, a visual comparison, an expert opinion, or measurements of an auxiliary variable related to the response. Then, based on these rankings, the response variable is measured for the ith ranked point from set i, with $i = 1, 2, ..., k$. Even though k^2 points are drawn, the response variable is only measured at k points. These k measurements represent a cycle, and m cycles are used to achieve a RSS of size $n = km$. Unbalanced RSS and RSS with random subsamples for $n < km$ are also possible (Amiri et al. 2015). This approach to data collection has led to an active field of research and many RSS approaches have been proposed (c.f. Wolfe 2010, 2012; Bouza-Herrera and Al-Omari 2019). Of particular interest for this article is the work by Robertson et al. (2021) on quasi-random RSS (QRSS).

QRSS uses a quasi-random number sequence, rather than a pseudo-random number sequence, to draw a ranked set sample of k locations from a continuous natural resource. To begin, the resource is first scaled and translated to fit within the unit box. Then, using a ranking set size k, the first k^2 points from a random-start Halton sequence (Wang and Hickernell 2000) that fall within the resource are chosen. These SB locations are divided into k distinct ranking sets, with each set containing k locations. Each ranking set is judgment ranked and based on these rankings, the response variable is measured for the ith ranked location from set i, with $i = 1, 2, ..., k$. Although the locations of the ranking variable are SB, there is no guarantee that the locations of the response variable are well spread. The authors showed that spreading the locations of the ranking variable in this way can substantially improve the precision of the sample mean when compared with balanced RSS.

In this article, we propose a new sampling strategy that incorporates local ranking information into a SB sample. The local ranking strategy we propose is simple:

1. Draw a SB sample of locations and at each location, judgment rank a set of nearby secondary sampling units (ssus); and

2. Measure the response variable at one judgment ranked ssu at each location.

The motivation for our approach is twofold. Firstly, to improve the precision of SB sampling by incorporating local ranking information into the design. Because the judgment ranked ssus are near the SB locations, a locally ranked sample is also SB. Hence, we expect the locally ranked approach to perform at least as well as the SB design used to draw the locations. Secondly, to provide a RSS approach that is not too onerous to implement in practice. Our method ensures that field technicians measure the response and ranking variable at

nearby sites to assist data collection. For example, ranking neighboring vines in vineyard, rather than ranking distant vines, before counting the number of fruiting buds (see Sect. 4).

The rest of this article is organized as follows. In Sect. 2, SB sampling with local ranking is introduced and estimation techniques are discussed in Sect. 2.2. Local ranking is numerically tested in Sect. 3 and an example application for the health and productivity assessment of a Shiraz vineyard in South Australia is given in Sect. 4. This assessment is integral to the management decisions on pruning or harvesting, and an improvement in sampling efficiency or accuracy enables better management actions. Concluding remarks are given in Sect. 5.

2. LOCALLY RANKED SAMPLES

To perform SB sampling with local ranking, a sampling frame consisting of primary sampling units (psus) and ssus is needed such that each psu contains nearby ssus. A sampling frame of this form can be constructed in a number of ways.

1. Merge nearby units/plots/panels into larger distinct groups.

2. Exploit naturally occurring structure in the population, for example, row/column or grid structure (see Sect. 4), blocks or panels.

3. Partition units/plots/panels into smaller ssus. For example, divide a large plot into two smaller plots to reduce the data collection effort at each location.

In this article, we assume a sampling frame containing spatial coordinates for N psus

$$\Omega_{\text{psu}} = \{x_1, ..., x_N\} \subset \mathbb{R}^2,$$

where x_i is the location of the ith psu with response value y_i and inclusion probability π_i. Furthermore, the ith psu is partitioned into k ssus $\{x_{i,1}, ..., x_{i,k}\}$ such that

$$y_i = \sum_{j=1}^{k} y_{i,j}, \tag{1}$$

where $y_{i,j}$ is the response at $x_{i,j}$ (the jth ssu at the ith psu). Other than satisfying (1), there are no restrictions on how the ssus are formed. Therefore, the ssus should be defined to assist data collection and make ranking simple.

2.1. SPATIALLY BALANCED SAMPLES WITH LOCAL RANKING

Consider estimating the population total

$$\tau = \sum_{i=1}^{N} y_i = \sum_{i=1}^{N} \sum_{j=1}^{k} y_{i,j},$$

by measuring the response variable at a sample of $n = km + \alpha$ ssus, where $k, m \geq 1$ and $0 \leq \alpha < k$ are integers. A locally ranked SB sample with ranking set size k is drawn using the following two-stage sampling approach.

1. Draw a SB sample of $n = km + \alpha$ psu locations from Ω_{psu}, given by

$$S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}. \tag{2}$$

2. Judgment rank the ssus at the sampled locations using a consistent ranking method for a measured variable

$$\begin{aligned} \{x_{1[1]}, \ x_{1[2]}, \ ..., \ x_{1[k]}\} \\ \{x_{2[1]}, \ x_{2[2]}, \ ..., \ x_{2[k]}\} \\ \vdots \qquad \vdots \qquad \qquad \vdots \\ \{x_{n[1]}, \ x_{n[2]}, \ ..., \ x_{n[k]}\}, \end{aligned} \tag{3}$$

where $x_{i[j]}$ is the jth ranked ssu at the ith psu location, $\boldsymbol{x}_i$.

3. Randomly select one ssu from each row of (3) such that

$$m \leq m_j \leq m + \lceil \alpha/k \rceil,$$

where m_j is the total number of jth ranked ssus in the sample with $\sum_{j=1}^{k} m_j = n$.

Any SB design can be used at step (1) to draw the sample locations and several designs are considered in Sect. 3. The ssus in (3) define a total of kn candidate ssus and the locally ranked sample contains one ssu from each row. Because each rank is included at least m but no more than $m + 1$ times, the inclusion probability of the ith psu, π_i, is spread evenly among its ssus. The inclusion probability of the jth ranked ssu at location $\boldsymbol{x}_i$, $x_{i[j]}$, is

$$\pi_{i[j]} = \pi_i \left(\frac{\alpha}{k} \left(\frac{m+1}{mk+\alpha} \right) + \left(1 - \frac{\alpha}{k} \right) \frac{m}{mk+\alpha} \right) = \pi_i/k \tag{4}$$

with $\sum_{i=1}^{N} \sum_{j=1}^{k} \pi_{i[j]} = n$.

There is flexibility in how the ssus are drawn at step (3). Provided the sample order in S is randomized, the locally ranked SB sample is

$$S_{\text{LR}} = \{x_{1[1]}, ..., x_{m_1[1]}, x_{m_1+1[2]}, ..., x_{m_1+m_2[2]}, ..., x_{n-m_n+1[k]}, ..., x_{n[k]}\}, \tag{5}$$

where m_j ssus of rank j are included in the sample. In this article, we force spatial spread in the locations for each rank in (5), such that

$$\{\boldsymbol{x}_1, ..., \boldsymbol{x}_{m_1}\}, \{\boldsymbol{x}_{m_1+1}, ..., \boldsymbol{x}_{m_1+m_2}\}, ..., \{\boldsymbol{x}_{n-m_n+1}, ..., \boldsymbol{x}_n\} \tag{6}$$

are SB subsamples from S. If GRTS or HIP is used at step (1), the sample output is ordered such that the sets in (6) are SB samples (Stevens and Olsen 2004; Robertson et al. 2018). Otherwise, the sets in (6) are obtained by sequentially using the SB design to draw equal

probability samples of size m_j from $\mathcal{S}$, removing the sampled locations from $\mathcal{S}$ after each iteration.

2.2. ESTIMATION

Locally ranked SB samples have first order inclusion probabilities (4), so the Horvitz–Thompson estimator

$$\widehat{\tau} = k \sum_{i=1}^{n} \frac{y_i}{\pi_i} \tag{7}$$

is an unbiased estimator τ, where y_i is the response value for $x_{i[j]} \in \mathcal{S}_{\text{LR}}$. Because the psus contain nearby ssus and the psus are SB, a locally ranked SB sample of ssus (regardless of the rankings) is spatially balanced. Therefore, the variance of $\widehat{\tau}$ depends explicitly on the SB design used to draw the psu locations. Forcing spatial spread in a sample can make second order inclusion probabilities difficult to calculate or close to zero (Stevens and Olsen 2003). For some SB designs, many second order inclusion probabilities are zero, meaning a design-based unbiased variance estimator is not possible (Grafström and Schelin 2014; Robertson et al. 2018).

Several variance estimators have been proposed for SB designs. The local mean variance estimator (Stevens and Olsen 2003) was proposed for GRTS and is commonly used, given by

$$\widehat{\text{var}}_{\text{NBH}}(\widehat{\tau}) = \sum_{i=1}^{n} \sum_{j \in D_i} w_{ij} \left(\frac{y_j}{\pi_j} - \widehat{\tau}_{D_i} \right)^2, \tag{8}$$

where D_i is a neighborhood containing at least four nearest neighbors from the sample to the ith point, $\widehat{\tau}_{D_i}$ is an estimate of the population total on D_i and w_{ij} are weights (see Stevens and Olsen 2003). Although this estimator does not explicitly include ranking information, each D_i will likely include different ranked ssus because the locations of each rank in $\mathcal{S}_{\text{LR}}$ are SB samples (see 6).

Another estimator, based on squared local deviations (Grafström and Schelin 2014) was proposed for LPM, given by

$$\widehat{\text{var}}_{\text{SB}}(\widehat{\tau}) = \sum_{i=1}^{n} \frac{n_i^*}{n_i^* - 1} \left(\frac{y_i}{\pi_i} - \frac{1}{n_i^*} \sum_{j \in S_i^*} \frac{y_j}{\pi_j} \right)^2, \tag{9}$$

where S_i is a subset of the sample containing n_i^* points. This subset includes all the sample points within a distance d of the ith point (including the ith point itself), where d is the distance to the nearest neighbor in the sample. Robertson et al. (2018) tested (9) and found it also worked well for HIP.

We also consider a resampling method proposed by Robertson et al. (2021) for equal probability quasi-random sampling, which is based on modified bootstrap RSS (Modarres et al. 2006). The variance estimator is given in Algorithm 1.

Algorithm 1 *Variance estimator for $\widehat{\tau}$ under locally ranked SB sampling,* $\widehat{\mathrm{var}}_{\mathrm{Boot}}$

1: **Input:** $\mathcal{S} = \{x_i, y_i\}_{i=1}^n$, where y_i is the measured response value at x_i.
2: **for** $b = 1$ to B **do**
3:     Define $\mathcal{S}_b^* = \{x_i^*, y_i^*\}_{i=1}^n$, where x_i^* is the ith point in a SB sample from
4:     Ω_{psu} and y_i^* is the response value of x_i^*'s nearest neighbour in $\mathcal{S}$.
5:     $\widehat{\tau}_b^* = \frac{N}{n} \sum_{i=1}^n y_i^*$.
6: **end for**
7: **Output:** $\widehat{\mathrm{var}}_{\mathrm{Boot}} = \frac{1}{B} \sum_{b=1}^B (\widehat{\tau}_b^* - \overline{\tau}^*)^2$, where $\overline{\tau}^* = \frac{1}{B} \sum_{b=1}^B \widehat{\tau}_b^*$.

At each pass of the outer loop (lines 2–6), an estimate of the population total $\widehat{\tau}_b^*$ is calculated from resampled response values. Lines 3 and 4 draw n SB psu locations from Ω_{psu} and for each psu, the nearest known response value from $\mathcal{S}$ is included in the resample. Resampling response values using nearest neighbors instead of random sampling with replacement aims to capture the spatial spread of locally ranked SB samples. Local ranking information is implicitly captured because the locations of each rank in $\mathcal{S}_{\mathrm{LR}}$ are SB samples. The outer loop gives B estimates of τ and the variance of these estimates is an estimator for $\mathrm{var}(\widehat{\tau})$.

We test the performance of $\widehat{\mathrm{var}}_{\mathrm{NBH}}$, $\widehat{\mathrm{var}}_{\mathrm{SB}}$ and $\widehat{\mathrm{var}}_{\mathrm{Boot}}$ for locally ranked GRTS, HIP, LPM and PWD samples in Sect. 3.3.

3. NUMERICAL RESULTS

In this section, we test the performance of equal probability SB sampling with local ranking on several synthetic populations. We restrict our attention to psus containing two ssus to investigate locally ranked samples in their simplest and most practical setting, where two neighboring units are ranked. The psu locations were defined using 100×100 regular grid over the unit box

$$\Omega_{\mathrm{psu}} = \{x_i = (2\lambda_1 + 1, 2\lambda_2 + 1)/200 : i = 1, 2, ..., 10,000\},$$

where λ_1, λ_2 are integers from $0, 1, ..., 99$. The locations of two neighboring psus at each x_i were defined using

$$x_{i,1} = x_i - (1/400, 0) \quad \text{and} \quad x_{i,2} = x_i + (1/400, 0)$$

so that

$$\Omega_{\mathrm{ssu}} = \{x_{i,j} : i = 1, 2, ..., 10,000 \text{ and } j = 1, 2\}.$$

We consider estimating the population total

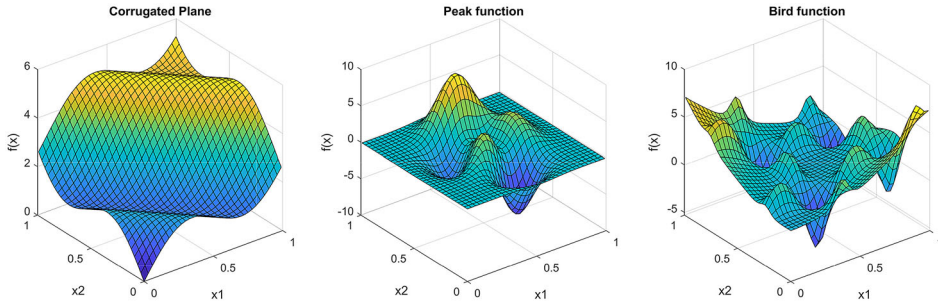$$\tau = \sum_{i=1}^{10,000} y_i = \sum_{i=1}^{10,000} \sum_{j=1,2} y_{i,j},$$

Figure 1. Three functions used to define the global spatial structure of the response variable.

where $y_i = y_{i,1} + y_{i,2}$ is the ith psu total and $y_{i,j}$ is the response value for $\boldsymbol{x}_{i,j}$. Although the spatial locations of the ssus are not needed to implement our local ranking strategy, defining them allows us to use existing designs that estimate τ by directly sampling either Ω_{psu} or Ω_{ssu}. Hence, the benefits of merging neighboring units together to define larger psus or dividing each psu into two smaller ssus can be investigated.

The response values for each population were defined via a continuous function f : $[0, 1)^2 \to \mathbb{R}$ such that

$$y_{i,j} = \alpha_{i,j} f(\boldsymbol{x}_{i,j}),$$

where $\alpha_{i,j}$ is a uniform random variable on $[1 - \ell, 1 + \ell]$ with $0 < \ell < 1/3$. The function f determines the global spatial structure of the response and the parameter ℓ removes the smoothness from f by introducing random jumps/drops to increase local variability in the response. We consider three functions from Robertson et al. (2018) and three levels of local variability with $\ell \in \{0.1, 0.2, 1/3\}$ to give a total of nine populations. These functions are illustrated in Fig. 1 and reader is referred to the supplementary material for further details.

The three ℓ values define different levels of local variability in the response. However, for each population, the local variability is not large enough to obscure the spatial trends defined by f. For $\ell = 0.1$, for example, we have locally similar response values

$$0.9 f(\boldsymbol{x}_{i,j}) \leq y_{i,j} \leq 1.1 f(\boldsymbol{x}_{i,j})$$

with $y_{i,j}$ at most 10% above or below $f(\boldsymbol{x}_{i,j})$. The intracluster correlation coefficients (ICCs) for these 10% populations were between 0.95 and 0.98, indicating that most of the variability occurred between the psus. Whereas, when $\ell = 1/3$, one of the psu's response values can be twice as large as the other, with ICCs between 0.63 to 0.83. We refer to these three scenarios as locally similar response values (10%); moderate local variability (20%); and relatively high local variability in the response (33%).

3.1. SAMPLING DESIGNS

Seven sampling designs were considered in the simulation study.

1. *SRS (ssu) and SRS (LR):* SRS on Ω_{ssu} and local ranked SRS.

2. *HIP (ssu), HIP (psu), and HIP (LR):* SB sampling using HIP on Ω_{ssu}, Ω_{psu} and with local ranking.

3. *GRTS (ssu), GRTS (psu), and GRTS (LR):* SB sampling using GRTS on Ω_{ssu}, Ω_{psu} and with local ranking (spsurvey, Dumelle et al. 2022).

4. *LPM (ssu), LPM (psu), and LPM (LR):* SB sampling using LPM on Ω_{ssu}, Ω_{psu} and with local ranking (BalancedSampling, Grafström and Lisic 2019).

5. *PWD (ssu), PWD (psu), and PWD (LR):* SB sampling using PWD on Ω_{ssu}, Ω_{psu} and with local ranking (Spbsampling, Pantalone et al. 2020).

6. *RSS:* RSS on Ω_{ssu} with $k = 2$ and $k = 5$.

7. *QRSS:* QRSS on Ω_{ssu} with $k = 2$ and $k = 5$ (see Sect. 3.1.1).

The (psu) and (ssu) sampling designs are single-stage methods and do not use ranking information. The (ssu) designs draw sample locations directly from Ω_{ssu}. The (psu) designs are cluster sampling methods that sample Ω_{psu} rather than Ω_{ssu} directly. To begin, a sample of $n/2$ SB psus was drawn. Then, the two response values from each sampled psu were selected to give a total of n response values.

The remaining designs are two-stage methods that utilize ranking information. RSS and QRSS sample Ω_{ssu} directly, without using the population's primary/secondary unit structure. The locally ranked (LR) designs use the primary/secondary unit structure to draw their sample. To begin, a sample of psus was drawn from Ω_{psu}. Then, the ssus within each sampled psu were ranked and the response was measured. We assumed that perfect rankings were possible, but other than for ranking purposes, the magnitude of the ranking variable was not utilized in the estimation. Although ranking errors are likely to occur in practice, this allowed us to study RSS, QRSS and local ranking in their ideal settings.

3.1.1. Quasi-Random RSS Using HIP

In this section, we extend the QRSS design of Robertson et al. (2021) to point resources. QRSS uses the random-start Halton sequence to ensure the locations of the ranking variable are well spread over a continuous resource. For point resources, we use HIP to draw SB sample locations for the ranking variable.

To draw a QRSS sample of $n = km$ locations from Ω_{ssu}, the following approach was used. To begin, an equal probability HIP sample of k^2m locations was drawn from Ω_{ssu}. These locations were then randomly assigned to km ranking sets, S_1, S_2, ..., S_{km}, with each set containing k locations. The QRSS sample was defined using

$$\mathcal{Q} = \{\boldsymbol{x}_{1[1]}, ..., \boldsymbol{x}_{m[1]}, \boldsymbol{x}_{m+1[2]}, ..., \boldsymbol{x}_{2m[2]}, ..., \boldsymbol{x}_{(k-1)m+1[k]}, ..., \boldsymbol{x}_{km[k]}\},$$

where $\boldsymbol{x}_{i[j]} \in S_i$ had the jth lowest judgment ranking in set S_i. The main difference between QRSS and HIP with local ranking is that the locations of the response values in QRSS are not necessarily spatially balanced, only the locations of the ranking variable are well spread. Locally ranked HIP samples are SB on the ssus because each ranking set is a psu containing nearby ssus and the sampled psus are well spread over the resource.

3.2. PRECISION

To investigate the precision of a sampling design, $\text{var}(\widehat{\tau})$ was estimated using the empirical variance estimator over $K = 10,000$ runs

$$\widehat{\text{var}}_{\text{SIM}}(\widehat{\tau}) = \frac{1}{K} \sum_{i=1}^{K} (\widehat{\tau}_i - \tau)^2, \tag{10}$$

where $\widehat{\tau}_i$ is the estimate of τ for the ith sample. We report relative precision by dividing (10) by the variance of the Horvitz–Thompson estimator under SRS on Ω_{ssu}, where values less than one indicate greater precision in the proposed design.

Results for GRTS, HIP, LPM and PWD are displayed in Fig. 2. SB sampling was an effective strategy here because each population has spatial trends that are not obscured by the local variability, with the strongest trends in the Corrugated Plane populations. As expected, the precision of SB sampling decreased as local variability in the response was increased. For the 10% populations, the ICCs were between 0.95 and 0.98, which made local ranking superfluous. Otherwise, for each SB design, local ranking consistently performed better than SB sampling on Ω_{ssu}, with PWD (LR) performing the best. SB sampling on Ω_{psu} with random (not ranked) ssu choice was also considered, but the results were indistinguishable from SB sampling on Ω_{ssu}. These results suggest that the precision of a SB design on populations with spatial trends can be improved using our local ranking strategy.

Results for SB cluster sampling are given in the supplementary material, as shown in Fig. 4. For each design, SB cluster sampling on Ω_{psu} was less effective than SB sampling on Ω_{ssu}. The ICCs were relatively high (between 0.63 and 0.95) and so sampling Ω_{ssu} directly was a better strategy. In some cases, SB cluster sampling performed worse than SRS on Ω_{ssu}.

Results for locally ranked and RSS approaches are displayed in Fig. 3. SRS with local ranking performed slightly better than SRS on the 33% populations, otherwise no meaningful gains in precision were observed. The RSS approaches were more precise than SRS, with QRSS performing the best. The only difference between these ranked set designs is that QRSS forces spatial spread in the locations of the ranking variable. Incorporating this spatial aspect into the design proved useful here because these populations have spatial trends. As expected, the larger ranking set size of $k = 5$ produced better results for both approaches. However, these designs require substantially more ranking information, making them more demanding to implement in practice. For most of the populations, QRSS with $k = 2$ performed better than RSS with $k = 5$.

HIP (LR), LPM (LR) and PWD (LR) performed better than QRSS on most of the populations considered, with QRSS ($k = 5$) performing better than HIP (LR), LPM (LR) on the Peak and Bird 33% populations. However, QRSS ($k = 5$) ranks $5n$ SB locations before the response is measured. Locally ranked samples only need to rank $2n$ ssus at n psu locations and the response is measured without leaving each psu's location. Hence, it may not be fair to compare QRSS ($k = 5$) with HIP (LR) and LPM (LR) directly, without considering the impact of ranking error and the cost of sampling a larger number of units in the ranking process. The important difference between QRSS and locally ranked SB samples is that
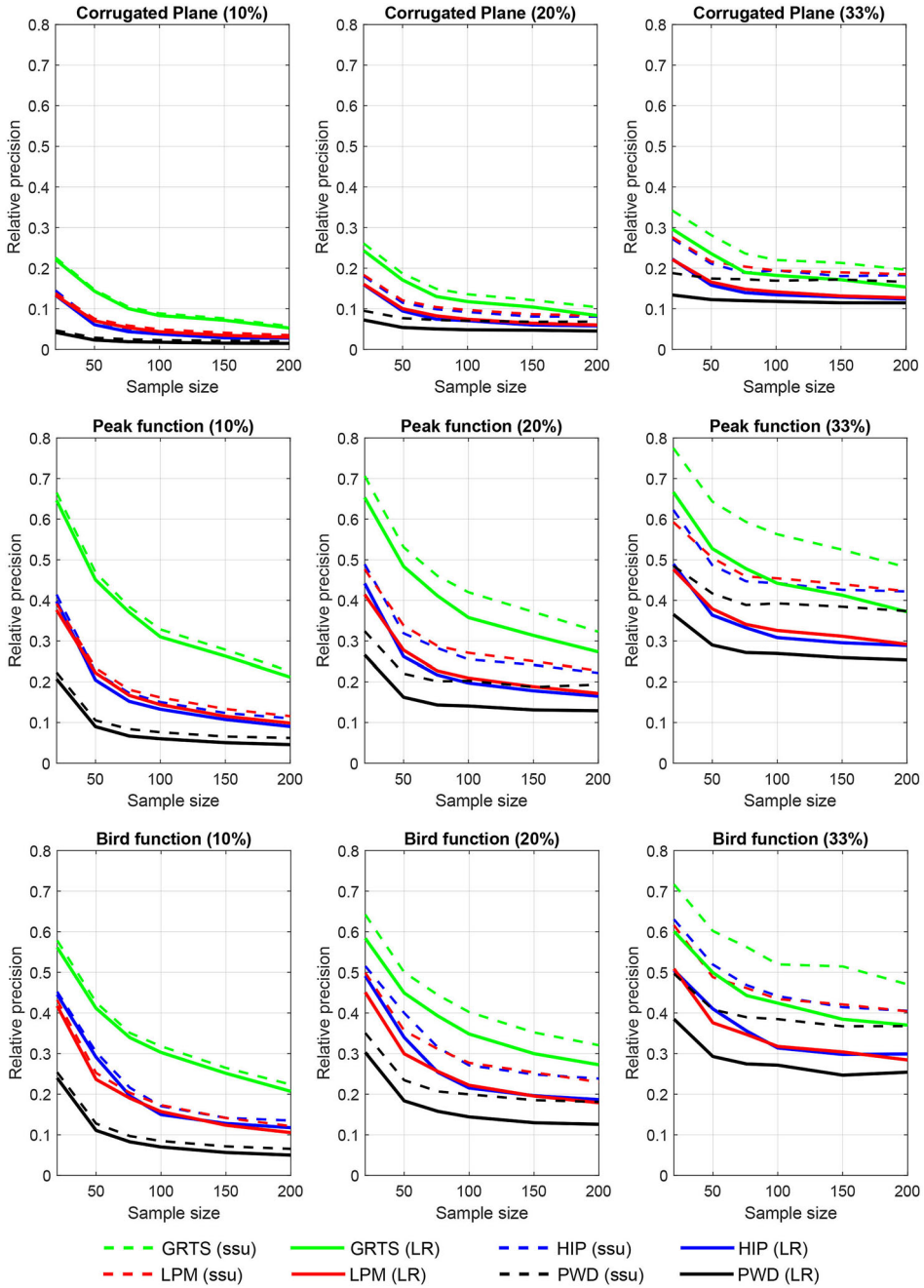
Figure 2.   Relative precision of GRTS, HIP, LPM and PWD on Ω_{ssu} and with local ranking on the Corrugated Plane, Peak and Bird functions with three levels of local variability in the response—locally similar (10%); moderate (20%); and relatively high (33%).

the locations of the response values in QRSS are not necessarily SB. Hence, forcing spatial spread in the locations of the ranking variable was an effective strategy, but retaining spatial spread in the locations of the response variable was also important.
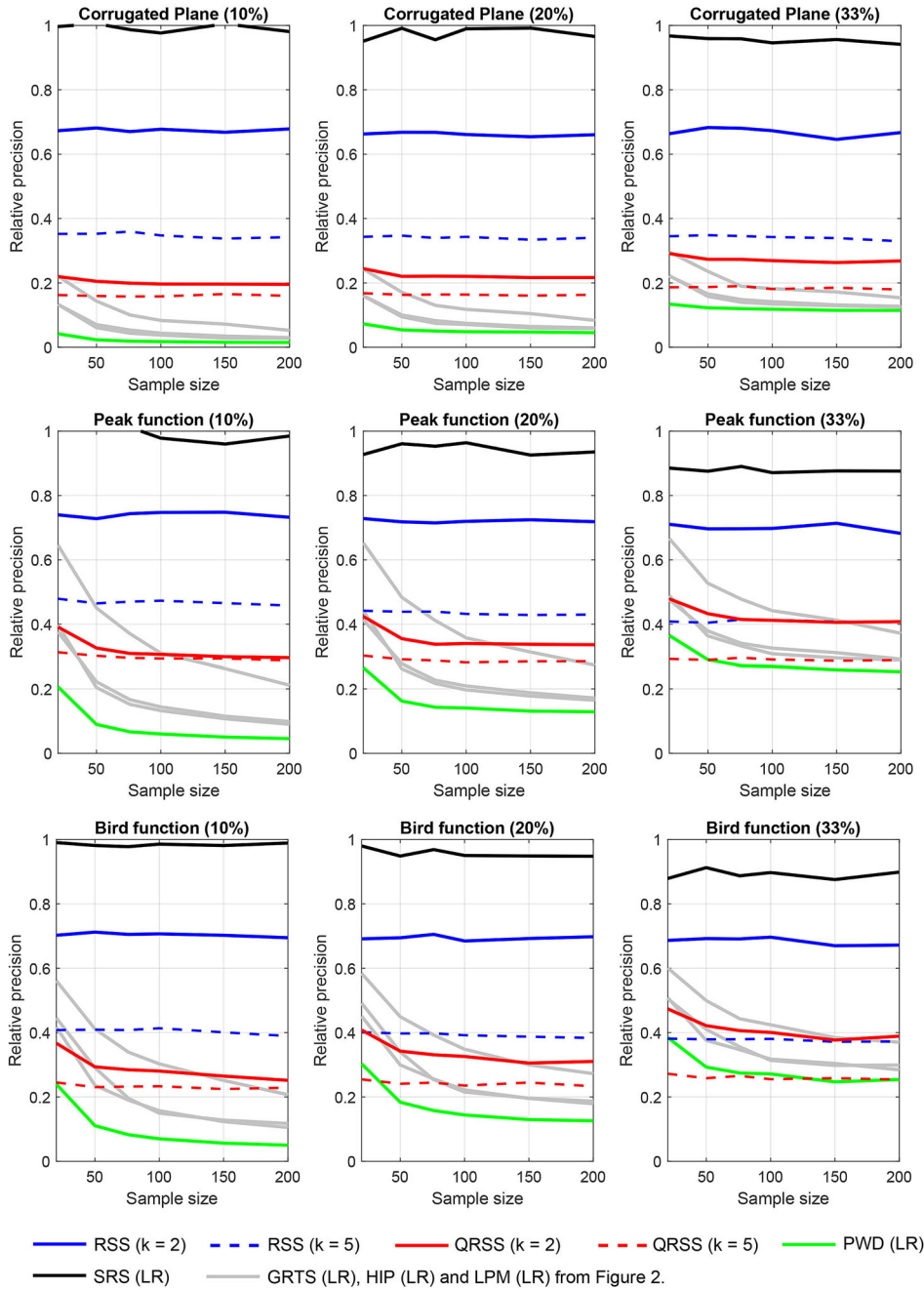
Figure 3. Relative precision of local ranking and RSS designs on the Corrugated Plane, Peak and Bird functions with three levels of local variability in the response—locally similar (10%); moderate (20%); and relatively high (33%).

From a practical point of view, the results for SB sampling with local ranking are encouraging. Firstly, with $k = 2$, ranking is not too cumbersome and ranking errors are less likely to occur. Secondly, because nearby ssus are ranked, we expect the implementation costs of

a locally ranked and unranked SB design to be similar. A more efficient approach (from a practical perspective) was SB cluster sampling, but the ICCs used here were too large to see statistical gains. We investigate the effect of low ICCs and weak spatial trends in Sect. 4. Finally, even if the local ranking is superfluous or unreliable, SB samples are obtained which are known to perform well on populations with spatial trends.

3.3. Variance Estimation

In this section, we consider estimating the variance of $\hat{\tau}$ for locally ranked SB samples using the nine populations from the previous section. We report the relative bias of the proposed estimator $\widehat{\text{var}}$ using $(\text{ave}(\widehat{\text{var}}) - \widehat{\text{var}}_{\text{SIM}})/\widehat{\text{var}}_{\text{SIM}}$, where $\widehat{\text{var}}_{\text{SIM}}$ is given in (10) and $\text{ave}(\cdot)$ is the average over 10,000 runs. The results are displayed in Fig. 4.

The variance estimators proposed for GRTS, HIP and LPM, performed reasonably well for locally ranked GRTS, HIP and LPM samples. The best estimator for locally ranked GRTS was the local mean variance estimator. To simplify Fig. 4, these curves were omitted because all relative biases for $\widehat{\text{var}}_{\text{NBH}}$ were between -0.05 and 0.15. The $\widehat{\text{var}}_{\text{NBH}}$ and $\widehat{\text{var}}_{\text{SB}}$ estimators provided conservative estimates for locally ranked LPM and HIP samples, with $\widehat{\text{var}}_{\text{SB}}$ having less bias on the 10% populations. The resampling variance estimator, $\widehat{\text{var}}_{\text{Boot}}$, also performed well for locally ranked LPM and HIP samples, and had the least relative bias for these methods.

The existing design-based variance estimators for SB approaches were not as effective for locally ranked PWD samples. The $\widehat{\text{var}}_{\text{NBH}}$ and $\widehat{\text{var}}_{\text{SB}}$ estimators were conservative and $\widehat{\text{var}}_{\text{Boot}}$ was a little optimistic on the 33% populations.

The best estimate of $\text{var}(\hat{\tau})$ for locally ranked SB samples depends explicitly on the SB design used to draw the sample. The authors recommend using $\widehat{\text{var}}_{\text{NBH}}$ for locally ranked GRTS samples and either $\widehat{\text{var}}_{\text{Boot}}$ or $\widehat{\text{var}}_{\text{SB}}$ for locally ranked HIP, LPM and PWD samples.

4. EXAMPLE APPLICATION: COOMBE VINEYARD

A vineyard rootstock trial is established at Coombe Vineyard, University of Adelaide, Waite Campus. The vineyard contains 350 Shiraz vines, which were planted in 1993 using a grid pattern with 11 rows and 32 columns. Each vine's trunk splits into two cordons (arms) of vines, which are trained on wire, and it is expected that thicker cordons are pruned to have more spurs (fruit bearing buds). See Johns (1957) for explanations and photos on training and pruning Shiraz plants. Estimating the total number of spurs can be useful for research and management purposes and can be incorporated into the overall assessment of the health and productivity of a vineyard. Therefore, an improvement in sampling efficiency or accuracy can enable better management actions. In 2019, University of Adelaide students conducted a census on 348 vines (two vines were not available) at the vineyard, where the number of spurs on each cordon was recorded. We use this population as an example application of SB sampling with local ranking to estimate the total number of spurs ($\tau = 5780$) at the vineyard. Exploiting the inherent vineyard structure, the vines are psus and the cordons are ssus, where each vine has two cordons to give a total of 696 cordons.
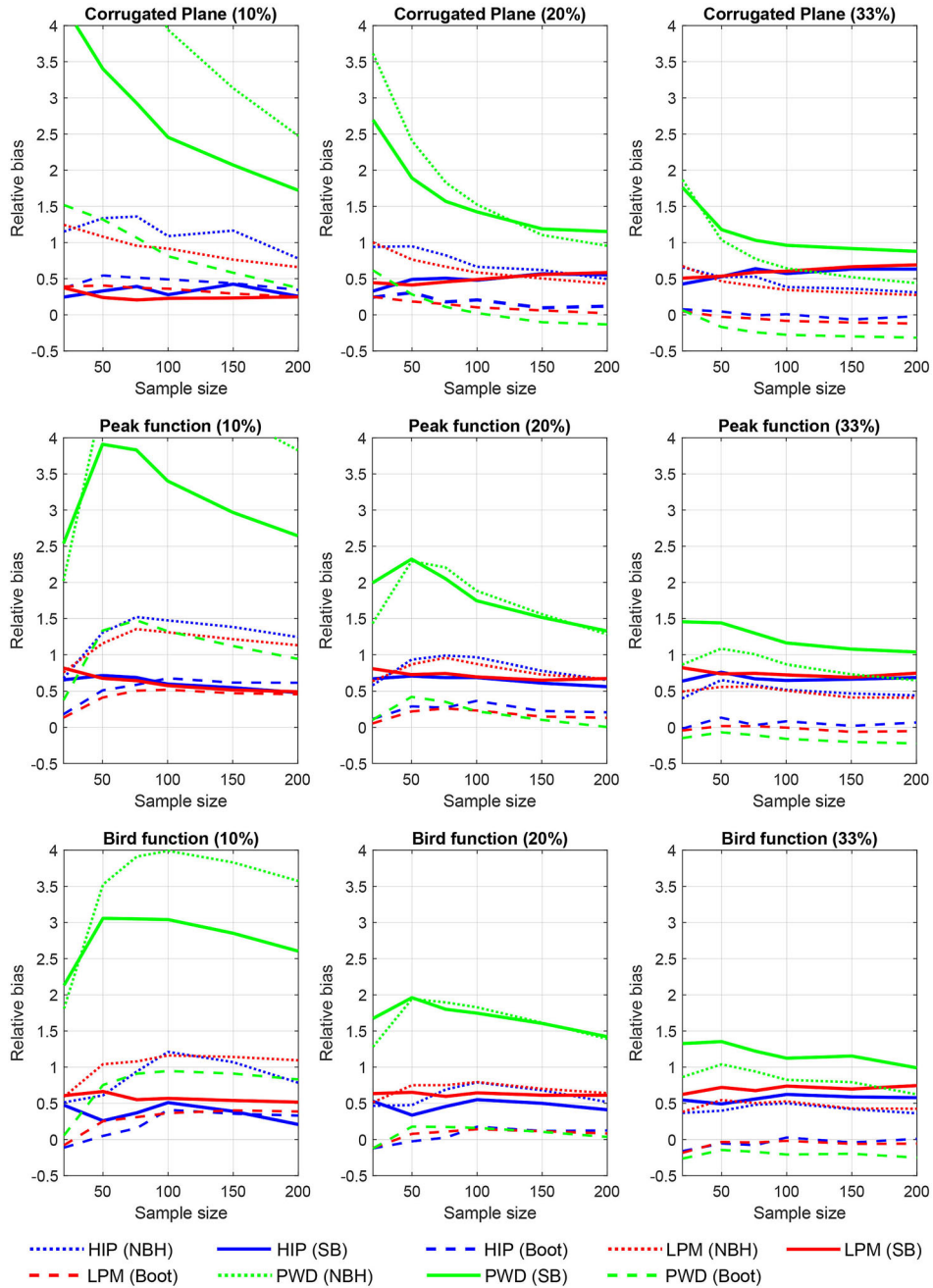
Figure 4. Relative bias of $\widehat{\mathrm{var}}_{\mathrm{NBH}}$, $\widehat{\mathrm{var}}_{\mathrm{SB}}$ and $\widehat{\mathrm{var}}_{\mathrm{Boot}}$ for SB sampling with local ranking on the Corrugated Plane, Peak and Bird functions with three levels of local variability in the response—locally similar (10%); moderate (20%); and relatively high (33%).

To draw a locally ranked SB sample, we begin by drawing a SB sample of $n = 2m + \alpha$ vines from the vineyard, where m is an integer and $\alpha \in \{0, 1\}$. Then, randomly choose $u \in \{0, 1\}$ and count the number of spurs on the lowest ranked cordons for $m + u$ vines

Table 1. Samples drawn from the Coombe Vineyard using four SB designs. Columns 2 to 4 show the relative empirical variance estimates of $\hat{\tau}$ for sampling the psus, ssus, and local ranking, respectively. Column 5 shows the empirical variance for local ranking and columns 6 to 8 show estimates of this variance. All values are averages using 10,000 runs

| | GRTS | | | | | | |
n	psu	ssu	LR	$\widehat{\mathrm{var}}_{\mathrm{SIM}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{NBH}}(\hat{\tau})$	–	–
20	0.96	0.99	0.67	99,404	113,200	–	–
50	0.95	0.99	0.66	37,530	46,203	–	–
100	0.94	1.00	0.61	16,010	23,107	–	–
150	0.94	0.98	0.58	9357	15,302	–	–
	HIP						
n	psu	ssu	LR	$\widehat{\mathrm{var}}_{\mathrm{SIM}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{NBH}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{SB}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{Boot}}(\hat{\tau})$
20	0.97	0.97	0.64	95,731	120,915	170,294	79,800
50	0.93	0.97	0.65	37,237	47,767	67,743	36,143
100	0.92	0.96	0.59	15,522	23,524	32,830	17,445
150	0.95	0.98	0.58	9328	15,594	21,875	9884
	LPM						
n	psu	ssu	LR	$\widehat{\mathrm{var}}_{\mathrm{SIM}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{NBH}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{SB}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{Boot}}(\hat{\tau})$
20	0.95	0.99	0.66	97,542	120,123	179,015	73,815
50	0.93	0.97	0.65	37,093	48,058	71,498	29,606
100	0.94	0.96	0.61	15,925	23,755	35,667	14,411
150	0.93	0.97	0.59	9403	15,710	23,614	8721
	PWD						
n	psu	ssu	LR	$\widehat{\mathrm{var}}_{\mathrm{SIM}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{NBH}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{SB}}(\hat{\tau})$	$\widehat{\mathrm{var}}_{\mathrm{Boot}}(\hat{\tau})$
20	0.95	0.97	0.65	97,308	123,543	186,237	52,328
50	0.93	0.97	0.63	35,903	48,714	73,163	22,661
100	0.94	0.97	0.60	15,834	24,093	36,861	12,359
150	0.90	0.96	0.57	9170	15,884	24,462	7918

in the sample. For the remaining $m + 1 - u$ vines in the sample, the spurs on the highest ranked cordons are counted. A desirable feature of this approach is that only n (rather than $2n$) locations need to be visited by the field technicians because the ranking is done locally at each vine.

To test the effectiveness of the local ranking strategy on this vineyard, we compare GRTS, HIP, LPM and PWD using local ranking and sampling the psus and ssus directly. Ranking errors were not permitted and we report relative precision by dividing each design's empirical variance (10) by $\mathrm{var}(\hat{\tau})$ under SRS. Two-stage SRS was also considered, but the results were indistinguishable from SRS. Results are given in Table 1.

Drawing SB samples of cordons (ssus) with GRTS, HIP, LPM and PWD produced results similar to SRS. Hence, spatial trends in the cordon counts were too weak to observe a statistical advantage from SB sampling. The ICC was low (-0.03) for this clustered population, and hence, drawing SB samples of vines (psus) was effective, with each SB design performing slightly better than SRS. However, the greatest improvements over SRS were made with local ranking.

Local ranking substantially improved the statistical performance of each SB design. SB sampling psus with random (not ranked) ssu choice was also considered, but the results were indistinguishable from SB sampling the ssus. In terms of practical implementation, the only

difference between a SB design and a locally ranked SB design is that the field technician needs to rank neighboring cordons before the spurs are counted. The $\widehat{\text{var}}_{\text{NBH}}$ estimator performed well for the locally ranked SB samples and $\widehat{\text{var}}_{\text{SB}}$ was overly conservative for HIP, LPM and PWD. The resampling estimator, $\widehat{\text{var}}_{\text{Boot}}$, produced good estimates for HIP, but consistently underestimated $\text{var}(\hat{\tau})$ for LPM and PWD.

5. CONCLUSION

In this article, we introduced a new sampling strategy that incorporates local ranking information into a SB sample. The strategy is simple and can be used with any SB design. To begin, a SB sample of psus containing nearby ssus was drawn. The ssus at each sampled psu were then judgment ranked, and based on these rankings, the response variable was measured at one ssu from each sampled psu. Because the psus contain nearby ssus, these samples are SB and not too onerous to select in practice.

Numerical results on test populations with different spatial structures and three levels of local variability showed that local ranking was an effective strategy for GRTS, HIP, LPM and PWD. If the response values within the psus were similar, local ranking was superfluous. However, if there was local variability in the response, local ranking improved the precision of the Horvitz–Thompson estimator. Existing design-based variance estimators for GRTS, HIP, LPM and PWD were tested and performed reasonably well for locally ranked samples on the populations considered.

Local ranking was also an attractive alternative to QRSS. Although QRSS's strategy of forcing spatial spread in the locations of a ranking variable was effective, retaining spatial spread in response variable was also advantageous. From a practical standpoint, local ranking has an advantage for field technicians because ranking nearby ssus assists data collection.

ACKNOWLEDGEMENTS

REFERENCES

Amiri S, Modarres R, Bhoj DS (2015) Ranked set sampling with random subsamples. J Stat Comput Simul 85:935–946

Benedetti R, Piersimoni F, Postiglione P (2015) Sampling balanced sampling: a review and a reappraisal. Int Stat Rev 85:439–454

Benedetti R, Piersimoni F (2017) A spatially balanced design with probability function proportional to within sample distance. Biom J 59:1067–1084

Bouza-Herrera CN, Al-Omari AIF (2019) Ranked set sampling, 65 years improving the accuracy in data gathering. Academic Press

Chauvet G, Le Gleut RH (2020) Inference under pivotal sampling: properties, variance estimation, and application to tesselation for spatial sampling. Scan J Stat, 48:1–24

van Dam-Bates P, Gansell O, Robertson BL (2018) Using balanced acceptance sampling as a master sample for environmental surveys. Methods Ecol Evol 9:1718–1726

Dickson MM, Tillé Y (2016) Ordered spatially balanced sampling by means of the traveling salesman problem. Comput Stat 31:1359–1372

Dumelle M, Kincaid TM, Olsen AR, Weber MH (2022) Spsurvey: spatial survey design and analysis. *R* package version 5.3.0

Grafström A (2011) Spatially correlated Poisson sampling. J Stat Plan Inference 142:36–41

Grafström A, Lundström NLP, Schelin L (2012) Spatially balanced sampling through the pivotal method. Biometrics 68:514–520

Grafström A, Lundström NLP (2013) Why well spread probability samples are balanced. Open J Stat 3:36–41

Grafström A, Tillé Y (2013) Doubly balanced sampling with spreading and restitution of auxiliary totals. Environmetrics 14:120–131

Grafström A, Schelin L (2014) How to select representative samples. Scand J Stat 41:277–290

Grafström A, Matei A (2018) Spatially balanced sampling of continuous resources. Scand J Stat 45:792–805

Grafström A, Lisic J (2019) BalancedSampling: balanced and spatially balanced sampling. *R* package version 1.5.5

Halls LK, Dale TR (1966) Trial of ranked set sampling for forage yields. Forest Sci 12:22–26

Halton JH (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer Math 2:84–90

Johns HK (1957) The pruning of grape vines. J Dep Agric 6:265–280

Kermorvant C, D'Amico F, Bru N, Caill-Milly N, Robertson B (2019) Spatially balanced sampling designs for environmental surveys. Environ Monit Assess 191:524

Larsen DP, Olsen AR Jr, Stevens DL (2008) Using a master sample to integrate stream monitoring programs. J Agric, Biol, Environ Stat 13:243–254

Modarres R, Hui TP, Zheng G (2006) Resampling methods for ranked set samples. Comput Stat Data Anal 51:1039–1050

McIntyre GA (1952) A method for unbiased selective sampling, using ranked sets. Aust J Agric Res 3:385–390

Pantalone F, Benedetti R, Piersimoni F (2020) Spbsampling: spatially balanced sampling. *R* package version 1.3.4

Robertson BL, Brown JA, McDonald T, Jaksons P (2013) BAS: balanced acceptance sampling of natural resources. Biometrics 3:776–784

Robertson BL, McDonald T, Price CJ, Brown JA (2017) A modification of balanced acceptance sampling. Statist Probab Lett 129:107–112

Robertson BL, McDonald T, Price CJ, Brown JA (2018) Halton Iterative Partitioning: spatially balanced sampling via partitioning. Environ Ecol Stat 25:305–323

Robertson BL, van Dam-Bate P, Gansell O (2021) Halton iterative partitioning master frames. Environ Ecol Stat 29:55–76

Robertson BL, Reale M, Price CJ, Brown JA (2021) Quasi-random ranked set sampling. Statist Probab Lett 171:109029

Stevens DL Jr, Olsen AR (1999) Spatially restricted surveys over time for aquatic resources. J Agric Biol Environ Stat 4:415–428

Stevens DL Jr, Olsen AR (2003) Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593–610

Stevens DL Jr, Olsen AR (2004) Spatially balanced sampling of natural resources. J Am Stat Assoc 99:262–278

Theobald DM, Stevens DL Jr, White DE, Urquhart NS, Olsen AR, Norman JB (2007) Using GIS to generate spatially balanced random survey designs for natural resource applications. Environ Manage 40:134–146

Wang X, Hickernell FJ (2000) Randomized Halton sequences. Math Comput Model 32:887–899

Wolfe DA (2010) Ranked set sampling. Comput Stat 2:460–466

Wolfe DA (2012) Ranked set sampling: its relevance and impact on statistical inference. ISRN Prob Stat 3:1–32