

Precision of systematic and random sampling in clustered populations: habitat patches and aggregating organisms

RICHARD MCGARVEY,^{1,3} PAUL BURCH,^{1,2} AND JANET M. MATTHEWS¹

¹*SARDI Aquatic Sciences, P.O. Box 120, Henley Beach, South Australia 5022 Australia*

Abstract. Natural populations of plants and animals spatially cluster because (1) suitable habitat is patchy, and (2) within suitable habitat, individuals aggregate further into clusters of higher density. We compare the precision of random and systematic field sampling survey designs under these two processes of species clustering. Second, we evaluate the performance of 13 estimators for the variance of the sample mean from a systematic survey. Replicated simulated surveys, as counts from 100 transects, allocated either randomly or systematically within the study region, were used to estimate population density in six spatial point populations including habitat patches and Matérn circular clustered aggregations of organisms, together and in combination. The standard one-start aligned systematic survey design, a uniform 10×10 grid of transects, was much more precise. Variances of the 10 000 replicated systematic survey mean densities were one-third to one-fifth of those from randomly allocated transects, implying transect sample sizes giving equivalent precision by random survey would need to be three to five times larger. Organisms being restricted to patches of habitat was alone sufficient to yield this precision advantage for the systematic design. But this improved precision for systematic sampling in clustered populations is underestimated by standard variance estimators used to compute confidence intervals. True variance for the survey sample mean was computed from the variance of 10 000 simulated survey mean estimates. Testing 10 published and three newly proposed variance estimators, the two variance estimators (v) that corrected for inter-transect correlation (v_8 and v_W) were the most accurate and also the most precise in clustered populations. These greatly outperformed the two “post-stratification” variance estimators (v_2 and v_3) that are now more commonly applied in systematic surveys. Similar variance estimator performance rankings were found with a second differently generated set of spatial point populations, v_8 and v_W again being the best performers in the longer-range autocorrelated populations. However, no systematic variance estimators tested were free from bias. On balance, systematic designs bring more narrow confidence intervals in clustered populations, while random designs permit unbiased estimates of (often wider) confidence interval. The search continues for better estimators of sampling variance for the systematic survey mean.

Key words: clustered populations; environmental monitoring; habitat patchiness; random sampling; spatial autocorrelation; survey precision; systematic sampling; transect; two-dimensional survey design; variance estimators

INTRODUCTION

One major challenge in field sampling is that most natural populations are not distributed uniformly or randomly in space (Perry et al. 2002). Plants and animals cluster for (at least) two reasons. First,

species are restricted to patches of suitable habitat, for example, in strata of favorable elevation, soil substrate, plant cover, or depth. Second, within patches of suitable habitat, individual organisms of the same species tend to aggregate, for reasons including predatory defense (Brönmark et al. 1984, Green and Nunez 1986, Dew 1990), enhanced reproductive success (Stevens et al. 1994, Babcock and Keesing 1999), or plant offspring settling near parents (Ehrlén and Eriksson 2000, Levine and Murrell 2003, Suzuki et al. 2005, Seidler and Plotkin 2006). Clustering renders field surveys of populations using quadrats or transects less precise, often yielding many zero counts

Manuscript received 15 October 2014; revised 26 March 2015; accepted 8 April 2015. Corresponding Editor: M. J. Ducey.

²Present address: Institute for Marine and Antarctic Studies (IMAS), University of Tasmania, c/o Australian Antarctic Division, 203 Channel Highway, Kingston, Tasmania 7050 Australia.

³E-mail: Richard.McGarvey@sa.gov.au

and a few very high counts. These skewed sample distributions greatly reduce the precision of survey estimates of mean population density (Pennington and Vølstad 1994, Pennington et al. 2002), or of quantities that depend on density. Clustered (spatially autocorrelated) populations pose a formidable cost challenge to environmental monitoring studies, requiring larger sample sizes to achieve equivalent levels of survey estimate precision. In this study we consider the two most common survey designs for environmental monitoring and ecological field study to assess their precision in populations that spatially cluster by these two processes.

The comparison of systematic and random survey methods was first active from the late 1930s to the 1950s. Statistical studies (Madow and Madow 1944, Cochran 1946, Yates 1948, Matérn 1960, Bellhouse 1977) concurred that systematic sampling schemes are more precise in autocorrelated populations. Some of this work (Madow and Madow 1944, Cochran 1946) was done assuming linear populations were to be sampled, such as an alphabetic list of households to be surveyed by telephone. A parallel succession of studies undertaken in the population biological literature (Bourdeau 1953), notably for application to sampling a two-dimensional space, surveying forests for timber yield estimation (Hasel 1938, Finney 1948, 1949), concluded that the precision advantage of systematic designs was minor and that random sampling is preferred (Greig-Smith 1983). More recent statistical investigation (D'Orazio 2003, Wolter 2007) assumed systematic sampling to be more precise in autocorrelated populations and focused on the still unsolved problem of how to reliably estimate the variance of the estimate of the mean from a systematic survey for linear (Wolter 1984, 2007) and two-dimensional (i.e., spatial) populations (D'Orazio 2003). Published studies in the applied ecological literature have now shifted to general agreement that systematic designs are more precise in spatially autocorrelated populations (e.g., Dunn and Harrison 1993, Ambrosio et al. 2004, Aune-Lundberg and Strand 2014).

In practice, field ecologists continue to apply both random and systematic sampling designs (Legendre et al. 2002). Random sampling satisfies the assumption of independence among samples, assuring accurate estimates of confidence interval by the standard error formula, $\sqrt{s^2/n}$, where the statistic s is the estimator of standard deviation and n is the number of observations. Systematic designs are often chosen for ease of design and implementation. Both random and systematic designs are unbiased (for systematic, see Madow and Madow [1944]), a major advantage achieved with no prior knowledge of the population to be sampled. Freedom from bias is not guaranteed with more sophisticated model-based or adaptive sampling approaches, though these can achieve higher precision when spatial model assumptions are met.

As a first objective, we evaluate and compare the precision of random and systematic survey designs for measuring population density in populations that cluster by habitat patchiness, aggregating behavior, or both. These two spatial processes, affecting probabilities or patterns of organism locations in space, are typical of many or most natural populations. High levels of clustering are a particularly common source of imprecision in marine surveys (Pennington 1996). Using simulated survey transects allocated randomly or systematically within a square two-dimensional study region, we quantified the imprecision of both survey designs by the variance of the repeated survey estimates, as the spread of replicated simulation survey mean densities. A measure of the cost advantage that may accrue from choosing the more precise design is inferred for each spatial population.

As a second objective, we evaluate variance estimators for the sample mean used to compute confidence intervals in systematic surveys. We tested 13 variance estimators, including 10 previously published (D'Orazio 2003, Wolter 2007). A comparison of 13 different variance estimators has not been undertaken, to our knowledge. Wolter's (1984, 2007) study, extended by D'Orazio (2003), has been the most comprehensive to date. We also tested three "covered grid" variance estimators proposed here (Appendix), which extend Yates' (1960) and Wolter's (2007) balanced difference method. In covered grid estimators, each balanced difference term uses all transects in each row or column of the systematic grid to uniformly cover the full width, the full length, or both, of the survey study region.

METHODS

Comparing the precision of random and systematic survey designs requires (1) simulated or enumerated spatial point populations of organism locations distributed within a study region, and (2) a method to simulate survey sampling of these spatial populations by the two survey designs to be tested.

Here, to test survey design performance in populations clustering by patchiness and aggregating organisms, we (1) generated six, and then a further 12, spatial point populations (Diggle 1983), each organism's location represented by an x - y coordinate point in a 1-km² study region, and (2) simulated transect sampling, measuring population density in each clustered (or unclustered) population. Using simulated rather than enumerated spatial populations permitted testing of a controlled range of clustering processes typical of natural populations, namely due to habitat patchiness and aggregating behavior, together and separately. Using populations with these combinations, we assessed the effect of these two forms of spatial clustering in natural populations for (1) comparing the relative precision of systematic vs. random sampling, and (2)

comparing 13 estimators for the variance of the mean from a systematic sample. Each simulated transect yields a single organism count, a sample measure of density. The accuracy and precision of each survey design was assessed by comparing replicated transect survey estimates of mean density with the true population density of each simulated population, which is known without error.

Simulated spatial populations

Six spatial point populations with differing degrees of patchiness and aggregating behavior alone or in combination were generated (Fig. 1). Population densities of the six simulated populations are given in Table 1 (row 1). Population a is spatially random, i.e., unclustered, with organism locations assigned independently and randomly. In populations b and c, organisms are again assigned randomly, but restricted to two drawn band-shaped patches of habitat covering ~22% of the study region. Populations d–f (Fig. 1d–f) are Matérn aggregated, in which organisms are distributed randomly within circles of 50 m radius, with the circle center points positioned randomly. The aggregations of population d are distributed across the entire study region (Fig. 1d), while populations e and f are Matérn aggregated within the two habitat patches (Fig. 1e, f). Matérn cluster circles can, by chance, randomly overlap (Fig. 1d–f), doubling the density in that aggregation. The two patches of habitat contain organisms of either approximately equal population density (Fig. 1b, e), or of densities that differed between the two patches by an approximate factor of 10 (Fig. 1c, f). These six simulated populations represent a wide range of clustering, from no clustering as complete spatial randomness (CSR, Fig. 1a; Diggle 1983) to both aggregated and patchy (Figs. 1e, f). The spatstat R package (Baddeley and Turner 2005) was used for both simulation steps of generating clustered populations and sampling them with transects (R code given in Supplement 1).

Simulated survey sampling: random and systematic

For each simulated survey, organism counts were obtained from a sample of 100 transects. Each transect was 1×50 m; 20 000 transects completely cover the study region without overlapping. These 20000 distinct transect positions provide the sample frame. For random surveys, 100 transects were selected independently and randomly, without replacement (Fig. 2, top right), from among the 20 000 possible transect locations. For systematic sampling, the 1-km² study region was partitioned into a 10×10 array of grid blocks, each a square of size 100×100 m (Fig. 2, left). Two rows of $100 \times 1 \times 50$ m transects cover each grid block. For each systematic sample of 100 transects, a first transect was chosen at random in the first (upper left) grid block. Following the standard procedure of one-start

(Madow and Madow 1944) aligned systematic sampling in a two-dimensional study region (Quenouille 1949, Cochran 1977), the randomly chosen start position within the first grid block was used for the position of all remaining transects, one transect in each of the other 99 grid blocks (Fig. 2, bottom right). Each survey yields a measured mean density as the mean from 100 transects. For each population and survey design, 10000 replicated Monte Carlo survey samples were drawn.

Accuracy and precision of the survey mean under random and systematic survey designs

We measure the accuracy (lack of bias) of each survey design by the closeness of the mean of 10000 replicated simulated survey density estimates to the true overall population density.

To compare the precision (inverse of variance) of the two survey designs, the true sampling variance of each survey design in each population was computed as the sample variance of mean density estimates from each set of 10000 simulated surveys. The cost advantage obtained by choosing the more precise survey design was inferred as the effective sample size of randomly allocated transects that would be required to equal the precision of the systematic survey design, knowing the variance of a randomly sampled mean varies inversely with sample size (as n in s^2/n).

Variance estimators for the survey mean

For systematic surveys, a variety of methods have been proposed to compute a confidence interval for the survey mean. In this study, we tested the performance only of design-based methods that use a single formula for the estimate of sampling variance. We estimate the accuracy and precision of each tested variance estimator for systematic surveys using simulation by comparing the (single) true variance (defined in the preceding subsection) with the mean of 10000 estimator variances, one for each survey of 100 simulated transects.

The variance estimator (v) of the mean for identical independent random sampling is s^2/n . Under sampling without replacement, the finite population correction $(1-f)$ (Cochran 1977) is included, giving what Wolter denotes as the textbook estimator $v_1 = (1-f)s^2/n$ (Wolter 2007:300). This correction is often small since f , the proportion of the population sampled, is small for most surveys. Following Cochran (1977) and Wolter (2007), the finite population correction was included in all 13 variance estimators tested.

Ten of the evaluated variance estimators for the systematic survey mean were from published sources. Wolter (2007) defined and assessed v_1 – v_8 . D’Orazio (2003) proposed v_{STR2} and v_W , modifying Wolter’s v_2 and v_8 to construct two-dimensional versions. In addition, we propose three new covered grid variance

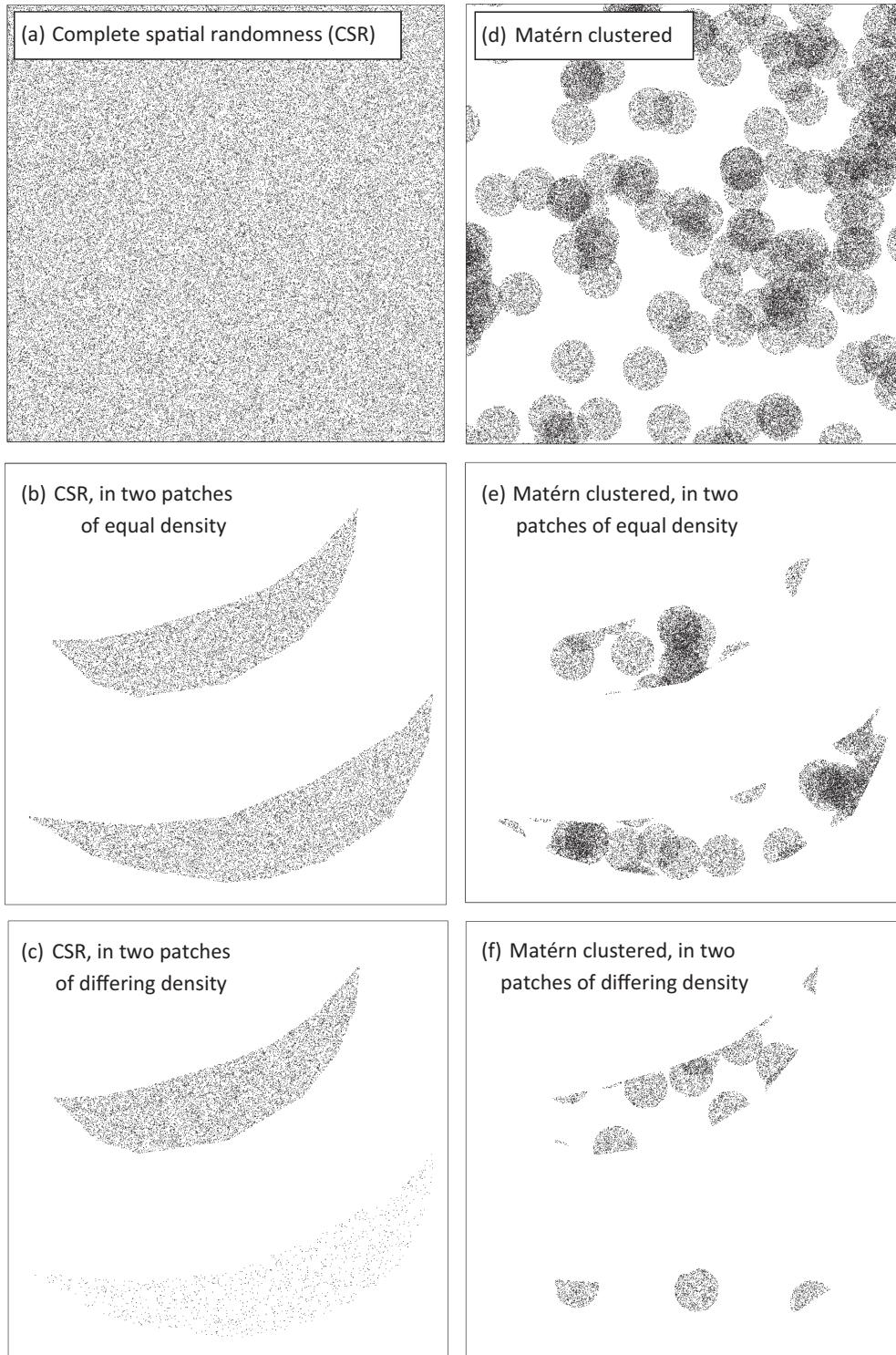


FIG. 1. Six simulated populations used for testing the performance of random and systematic survey designs. These populations express two forms of clustering: patchy habitat and aggregating behavior, here simulated by a Matérn process. Population a is generated as a complete spatial random (CSR) process. Populations b and c are CSR within two band-shaped patches, where in population c the lower patch has a density $\sim 1/10$ th smaller. Population d expresses only Matérn clustering with no patches, while organisms in e and f Matérn cluster within two habitat patches, in population f with patch densities again differing by factor of 10.

estimators (Appendix), extending the balanced difference method of Yates (1960) and Wolter (2007). In addition to the three covered grid estimators described in the Appendix, formulas are included for v_1 (the textbook estimator) and v_8 (given in *Discussion*). The remaining eight variance estimator formulas are detailed by their authors (D'Orazio 2003, Wolter 2007). The 13 variance estimators for systematic samples were coded as subroutines in R (R Core Team 2013; Supplement 2) and applied to each simulated systematic survey sample of 100 transects. Applying Wolter's v_2 - v_8 requires a specific ordering of transects, which was done here in a vertical orientation, starting from the top left transect, down each column of grid blocks, columns succeeding from left to right.

RESULTS

Bias of random and systematic survey designs

The mean densities of 10000 simulated surveys, for both simple random sampling (Table 1, row 2) and systematic sampling (Table 1, row 3), were close to the true population density (values in these two rows are close to 1), implying both survey designs are unbiased. Because these designs are both representative, neither using a priori information about population density in allocating samples within the study region, this result of non-bias is expected and well known. It was first proven for systematic sampling by Madow and Madow (1944).

True precision of random and systematic survey designs

While both random and systematic designs are unbiased, they differ in sampling variability precision. For all five clustered populations (b-f, Fig. 1b-f), with $n = 100$ transects, the systematic survey design was much more precise. The true variances for the systematic survey estimate of mean density (from 10000 survey estimates) were one-third (36%) to one-fifth (18%) of the true variances achieved by the random design (Table 1, row 5, columns b-f).

For population d, with the population clustering only by Matérn aggregating of organisms (Fig. 1d), true systematic survey variances were one-third (0.33, Table 1, row 5, column d) those of randomly allocated transects. Theory of effective sample size (Pennington and Vølstad 1994), based on the inverse linear decline of sampling variance with sample size under independent random sampling, implies that a survey using the random design would require three times as many transects (301 vs. 100, Table 1, row 6) to equal systematic precision, given the observation that the random sampling variance is three times wider.

When organisms were distributed with complete spatial randomness (no aggregating behavior), but were restricted in habitat range to two patches of either

roughly equal density (Fig. 1b), or density that differed by 10-fold (Fig. 1c), the systematic survey design again yielded true sampling variances that were much smaller, 29% and 23%, respectively, of those obtained using the random design (Table 1, row 5, columns b,c), implying that for these two populations that expressed clustering only via habitat patchiness, with 100 systematic transects, 3.4 and 4.3 times as many random transects (344 and 430 vs. 100, Table 1, row 6, columns b,c) would be needed to equal the precision obtained using the 10×10 square-grid systematic design.

When individuals were both Matérn aggregated and restricted to suitable habitat, systematic sampling variances for populations distributed in two patches of roughly equal density (Fig. 1e) or 10-fold differing densities (Fig. 1f) were 18% and 36% those of the 10 000 random surveys (Table 1, row 5, columns e,f), respectively, implying that over five or nearly three times (565 and 275 vs. 100, Table 1, row 6, columns e,f) as many random transects would be needed to equal systematic survey precision.

In summary, for these simulated clustered populations, a systematic 10×10 square grid of transects produced estimates of absolute population density that were three to five times more precise, thus requiring approximately one-third to one-fifth as many transect sample counts to achieve equivalent precision in the estimate of the sample mean. This precision advantage for systematic survey held about equally for the two tested sources of clustering in natural populations, habitat patchiness and aggregating behavior.

By contrast, in the purely spatially random (CSR) population (Fig. 1a), the systematic survey design was slightly less precise, the true variance of 10000 systematic survey estimates being 14% larger (1.14, Table 1, row 5, column a) than that obtained by random sampling.

Testing variance estimators for systematic surveys

For the random survey design, in all six populations, the textbook (sample) variance estimator (v_1) gave accurate agreement with the true variance (results not shown). This was expected because random survey designs satisfy the assumptions of identical independent sampling.

However under systematic sampling, the v_1 -estimated variances differed from the true variances in all six tested populations (Table 1, row 9), implying bias. In the CSR population, v_1 slightly underestimated (0.87) the true spread of systematic mean density estimates. In the five clustered populations, v_1 overestimated the true sampling variance, by 2.79 up to 5.63 times (Table 1, row 9, columns b-f), implying v_1 underestimated the true precision of systematic survey estimates of mean density in clustered populations.

The true variances of the systematic sample for populations a-f, each computed as the variance of

TABLE 1. Test statistics from simulation sampling by 100 transects in the six simulated populations of Fig. 1a–f.

Row	Test statistic	Non aggregating organisms			Matern aggregating organisms		
		a) No patches (CSR)	b) Two habitat patches of equal density	c) Two habitat patches of differing densities	d) No patches	e) Two habitat patches of equal density	f) Two habitat patches of differing densities
Section 1: testing for accuracy of random and systematic sampling							
1	True population density	0.100	0.022	0.011	0.105	0.025	0.007
2	(Mean density estimate)/(true density), random sampling	0.9989	0.9965	1.0046	1.0012	1.0008	0.9973
3	(Mean density estimate)/(true density), systematic sampling	0.9996	0.9996	1.0017	0.9997	1.0003	0.9950
Section 2: survey precision, random and systematic designs							
4	Variance of 10000 systematic survey mean density estimates (giving true sampling error variance)	2.28×10^{-5}	5.29×10^{-6}	2.27×10^{-6}	4.53×10^{-5}	1.05×10^{-5}	2.87×10^{-6}
5	Relative precision: ratio of true sampling variances, systematic over random	1.14	0.29	0.23	0.33	0.18	0.36
6	Effective sample size of random transects needed to equal true variance observed for 100 systematic transects	88	344	430	301	565	275
7	Percentage of zero-count systematic transects	1%	74%	83%	34%	83%	91%
Section 3: testing variance estimators of Wolter (2007) for a systematic sample mean							
8	v_1 variance estimate (textbook SE^2)	2.11×10^{-5}	1.84×10^{-5}	2.14×10^{-5}	2.21×10^{-5}	2.19×10^{-5}	1.84×10^{-5}
9	v_1 /(true variance): mean of v_1 -estimated variances (from 10000 systematic sample replicates) divided by true variance	0.87	3.52	4.38	3.12	5.63	2.79
10	v_2 /(true variance)	0.88	3.03	3.43	2.65	4.98	2.74
11	v_3 /(true variance)	0.89	3.13	3.88	2.72	5.16	2.80
12	v_4 /(true variance)	0.88	2.59	2.97	2.53	4.62	2.69
13	v_5 /(true variance)	0.88	2.21	2.48	2.57	4.33	2.67
14	v_6 /(true variance)	0.89	2.26	2.41	2.88	4.23	2.50
15	v_7 /(true variance)	0.87	4.75	3.10	1.18	4.55	1.06
16	v_8 /(true variance)	0.66	1.14	1.13	1.09	2.19	2.06
Section 4: testing two D'Orazio (2003) and three proposed (covered grid) variance estimators of the systematic sample mean designed for use in a two-dimensional study region							
17	v_{STR2} /(true variance)	0.88	2.52	3.09	2.73	5.04	2.81
18	v_W /(true variance)	0.68	0.62	0.7	1.19	1.68	1.39
19	v_{RCG} /(true variance)	0.88	1.19	1.59	2.44	3.73	2.05
20	v_{CCG} /(true variance)	0.90	2.02	2.02	2.81	3.93	1.69
21	v_{TCG} /(true variance)	0.90	1.61	1.81	2.62	3.83	1.87

Notes: Section 1 assesses bias of random and systematic sampling, giving the means of 10000 survey-measured densities compared with true population densities. Section 2 compares true precision of random with systematic designs. Sections 3 and 4 show bias of 13 tested variance estimators under systematic sampling, quantified by the means of 10000 simulation-survey estimated sampling error variances (v) divided by the true sampling error variance for each population. Covered grid variance estimators are row-covered (RCG), column-covered (CCG), and twice-covered (TCG).

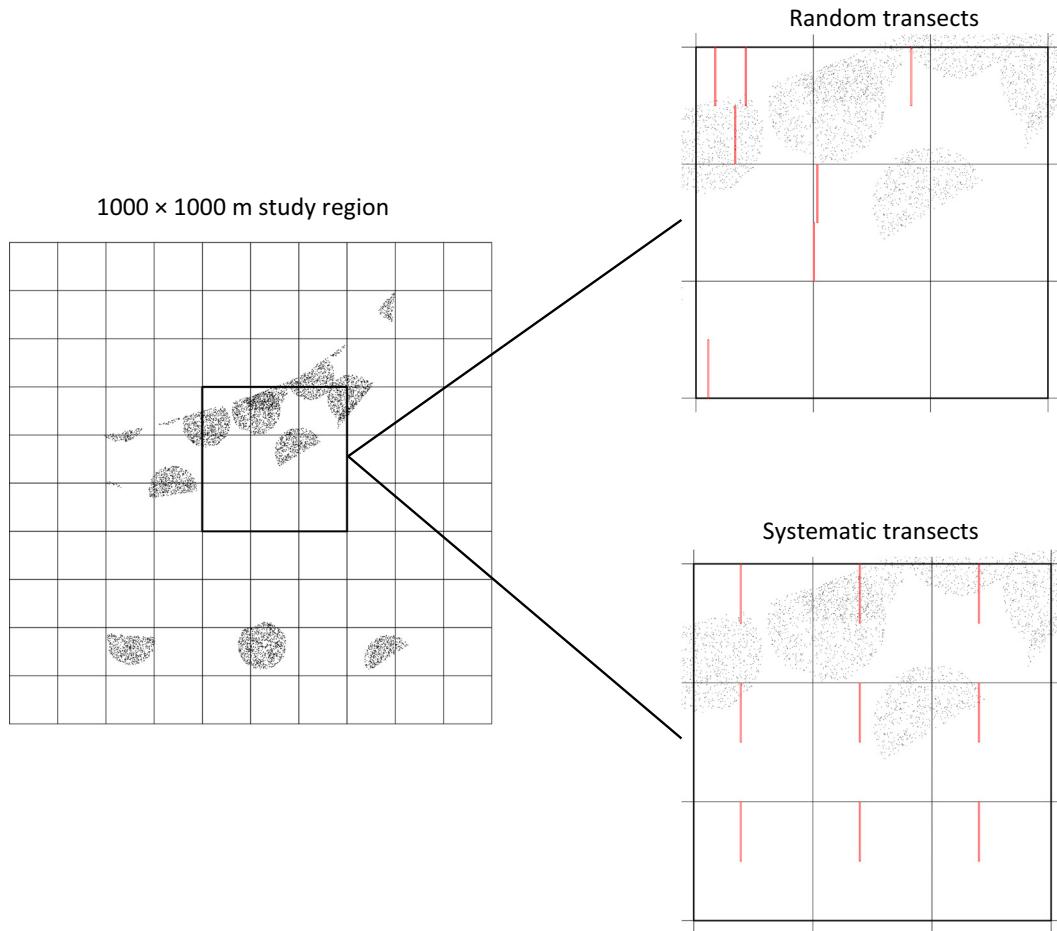


FIG. 2. Diagram of the two transect sampling designs, random and systematic. For both designs, a simulated sample of 100 transects was allocated within the study region to measure absolute population density. Here, the population is Matérn clustered in patches of unequal density (Fig. 1f). The 1000 × 1000 m study region (left) was divided into 100 grid blocks, each 100 × 100 m, to construct the systematic design. Boxes to the right illustrate possible differences in transect placement for a 300 × 300 m central portion of the study region. Under (aligned) systematic sampling, transects (red lines) are uniformly spread over the study area (one transect allocated to the same position in each grid block). Under simple random sampling, each transect is allocated within the study region independently of others, resulting by chance in some grid blocks being sampled multiple times while others are left unsampled.

the replicated sample means from 10 000 simulated systematic surveys, are given in row 4 of Table 1. The accuracy of each variance estimator was quantified by the ratio of the mean of 10 000 variance estimates divided by the true variance. Mean values near 1 in Sections 3 and 4 of Table 1 imply accurate (unbiased) estimates of systematic sampling variance, and are graphically shown as red dots (means of 10 000 estimated variances) near the blue line (true variance) in Fig. 3. Values greater than 1 imply overestimated (overly wide) confidence intervals.

In the purely random CSR population (Fig. 1a), all 13 variance estimators underestimated the true variance of systematic sampling (Fig. 3a). Estimators other than v_8 and v_W underestimated by 10–13% (Table 1, column a, rows 9–21). The least biased were two of the covered grid estimators (0.90 for v_{TCG} [twice-covered

grid estimator] and v_{CCG} [column-covered grid estimator]); v_8 and v_W underestimated the true variance by about one-third.

One variance estimator, v_7 , performed particularly poorly, giving unacceptably wide ranges of variance estimates, with values reaching close to zero for all populations (Fig. 3) other than b (Fig. 3b), where it was highly biased. Means of v_7 well above medians (Fig. 3) imply large upper tails. This exceptionally unstable performance of v_7 is sufficient to rule it out for general use, as Wolter (2007) had earlier concluded for linear populations. We do not consider v_7 henceforth.

For the population of Fig. 1b, nonaggregating (CSR) populations restricted to two habitat patches of equal density, v_8 (1.14) and v_{RCG} (1.19; row-covered grid estimator) were the least biased (Table 1, column b), v_8 being the more accurate (Fig. 3b); v_W was

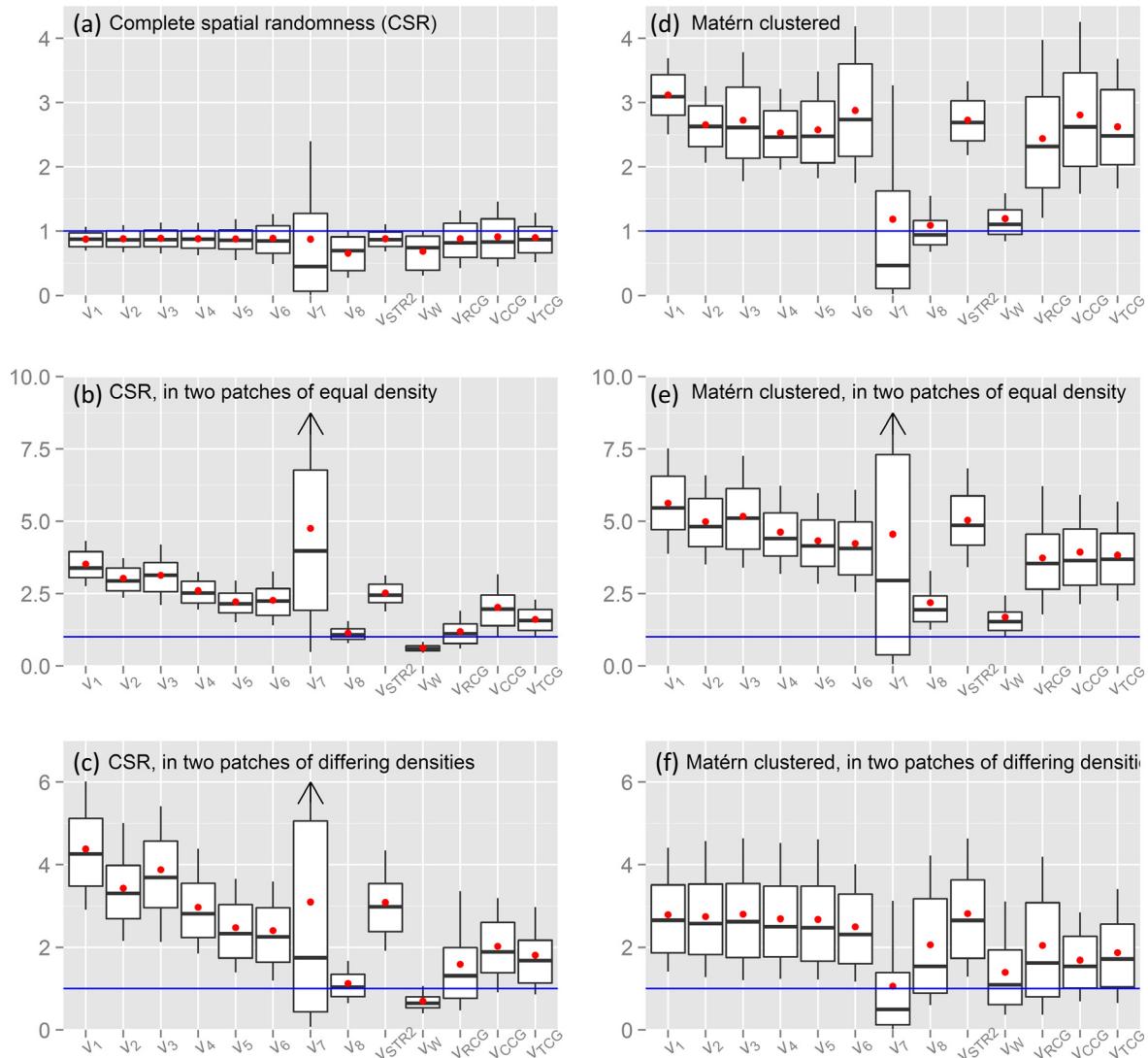


FIG. 3. Boxplots displaying systematic variance estimator precision and accuracy. Thirteen variance estimators (v_1 – v_{TCG}) for survey mean density were tested using replicated systematic sampling of the six populations of Fig. 1. The first 10 variance estimators include the eight presented by Wolter (2007), v_1 – v_8 , plus two extensions of Wolter estimators proposed by D’Orazio (2003), v_{STR2} and v_W . The three covered grid variance estimators (v_{RCG} , v_{CCG} , v_{TCG} ; row-covered, column-covered, and twice-covered, respectively) we propose here and describe in the Appendix. Each survey gave a measured mean density from 100 simulation transect counts. Box (25% and 75% quantiles) and whiskers (10% and 90% quantiles; arrows indicate that whiskers extend beyond the displayed y-axis range) display variance estimate precision as the spread of 10000 simulated estimated variances. The horizontal line in each box shows the median. Each red dot marker shows the mean of 10000 variance estimates. Y-axis scaling is given by dividing each estimator variance by the (single) true variance of 10000 replicate survey mean densities. Blue lines show the value (1) at which estimated survey variance agrees with true.

underestimation biased (0.62, Table 1, row 18, column b; Fig. 3b). Populations b and c showed similar results overall (Fig. 3b and c), with v_8 the best performer, being both accurate and precise, and with most others overestimating or greatly overestimating the true variance of the systematic sample mean; v_W considerably underestimated the true variance in these two populations. After v_8 , covered grid estimators were the next least overestimation biased (Fig. 3b and c).

In population d, with Matérn clustering but no patchiness of habitat (Fig. 1d), v_8 (1.09) and v_W (1.19) showed

relatively low bias (Table 1, column d, rows 16 and 18; Fig. 3d). The others (v_1 – v_6 , v_{STR2} , and three covered grid estimators, Table 1, column d; Fig. 3d) showed much higher and similar bias, overestimating the true variance on average by ~ 2.4 – 3.1 times. In addition to being nearly unbiased, v_8 and v_W also had much tighter spreads of variance estimates (boxes and whiskers, Fig. 3d), and so were also more precise.

Both patchiness and aggregation within patches were simulated in populations e and f. For population e (Fig. 1e) with Matérn clustering in two patches of

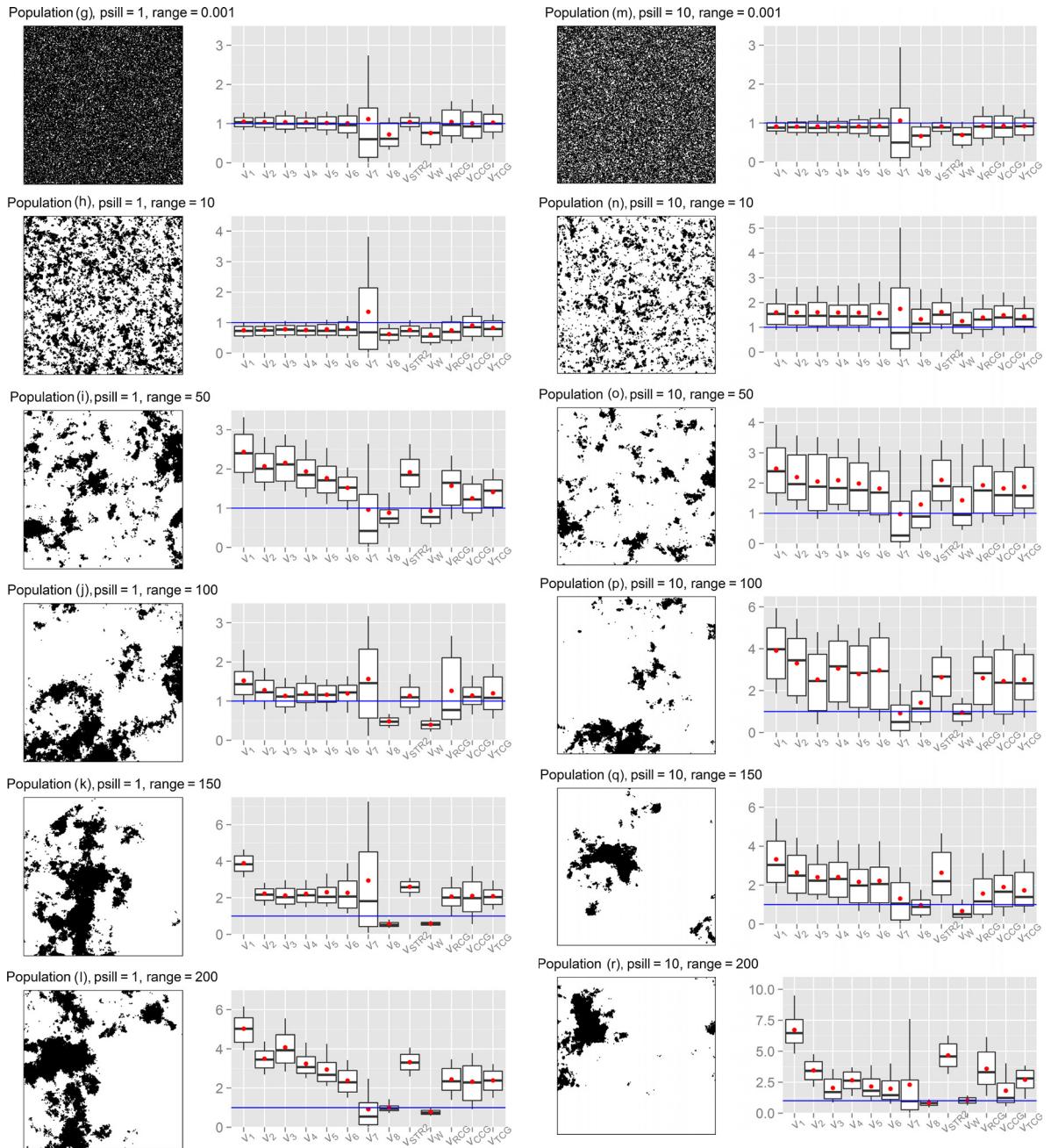


FIG. 4. Twelve spatially autocorrelated populations, generated using the R package *gstat*, where the degree of autocorrelation was controlled by two parameters of the variogram, *psill* and *range*. Boxplots show the spread of 10000 variance estimates for each of the 13 tested variance estimators, as in Fig. 3.

similar density, v_W (1.68) and v_8 (2.19) were again by far the least biased (Table 1, column e; Fig. 3e), though all overestimated. After v_W and v_8 , covered grid estimators showed the least, though still large, bias (~3.73–3.93, Table 1, column e). For this common ecological case of a population aggregating within habitat patches, v_8 and the closely related v_W were again clearly superior, showing both lower bias and

narrower interquartile ranges (Fig. 3e), though overestimation remained.

For the most highly clustered population (Fig. 1f), with Matérn clustering and restriction to two habitat patches of unequal density, all variance estimators performed poorly, overestimating and showing a wide spread of estimates; v_W showed the lowest bias (Table 1, column f, row 18) and a modestly smaller than

average interquartile range (Fig. 3f), while v_8 performed less well (Fig. 3f). Covered grid estimators were the next least biased (Table 1, column f).

Variance estimator comparison with a second set of spatial populations

The principal new result obtained from these simulations was the appearance of v_8 and v_W as more reliable estimators of sampling variance for the systematic survey mean in clustered populations. This was observed for the five spatial point populations (b–f), which express clustering by drawn habitat patch boundaries and Matérn (circular) aggregations. The question naturally arises, how do these results extend to other plausible clustered populations? To test for sensitivity to the spatial point populations generated, we used a different set of algorithms to produce a second set of 12 spatially autocorrelated populations (Fig. 4, populations g–r). To sample these new populations, the same simulated systematic survey design was applied (Fig. 2) using $n = 100$ 1×50 m transects. The 13 variance estimators were again applied to 10000 simulated systematic surveys in each population to evaluate estimator performance.

The autocorrelation properties of these new populations were controlled by two parameters of the variogram, *psill* and *range*. Following Goslee (2006), we used the *gstat* R package (Pebesma and Wesseling 1998, Pebesma 2004), which provides routines for kriging and simulation, to allocate organism point locations. The *range* specifies the separation distance beyond which points are effectively uncorrelated. For all simulated populations of Fig. 4, we set the nugget (i.e., measurement error of repeated sampling at the same location) equal to zero. With a zero nugget, the *psill* parameter we control equals the sill as normally defined, being the asymptotic value of the semivariance at separation distances greater than the *range*. Two values of *psill* (1 and 10) were chosen, and for each of these, six values of the *range* of autocorrelation: 0.001, 10, 50, 100, 150, and 200 m. The R code used to generate this second set of populations is given in Supplement 3.

By testing the two values of *psill*, we examine two suites of populations where the maximum density of tight clusters differ by an order of magnitude, broadening these results to apply for a wider range of possible natural field populations. Simulating a broad array of autocorrelation distances (*range* values) tests the robustness of these results from populations that are effectively uncorrelated in space (*range* = 0.001 m) to those correlated over long distances (*range* = 200 m) greater than the length of transects (50 m) and the distance between neighboring transects (100 m).

By this method of generating autocorrelated spatial populations, unoccupied areas arose naturally within the study region (see point maps of these 12

populations in Fig. 4). Clustering is visually evident in the point maps as *range* increases.

The results for this second set of 12 populations (Fig. 4) show qualitative similarity, with some differences, to those observed for the six populations described previously. The textbook estimator (v_1) was again consistently the worst performer, overestimating the sampling variance greatly for increasingly autocorrelated populations (Fig. 4 boxplots, populations i–l and o–r).

In Fig. 3, for the four populations restricted to patches of habitat (b, c, e, less strongly for f), the variance estimators can be approximately ranked in terms of bias starting with the worst bias of $v_1 > v_2 - v_3 > v_4 - v_6 > v_{RCG} - v_{TCG} > \text{or } \gg v_8$ and v_W . In Fig. 3, only the population with no habitat patches (d) did not express this trend, yielding similar overestimation bias for estimators $v_2 - v_6$ and $v_{RCG} - v_{TCG}$, though v_1 was again the worst, and v_8 and v_W were by far the best.

This same bias trend was evident for a majority of the *gstat* populations in Fig. 4, strongly (i, l, q) or weakly (j, o, p), but was less evident for others (k, r).

Again v_7 showed unacceptably poor performance for all 12 populations of Fig. 4. D’Orazio’s extension of v_2 , v_{STR2} , as in Fig. 3, showed high overestimation bias similar to v_2 , overestimating more than other estimators (except v_1) for the high-*range* autocorrelated populations (k, l, q, r).

The performance outcomes for v_8 and v_W were similar to those of Fig. 3, but differed between the two levels of *psill*. For *psill* = 10, v_8 and v_W , in the more clustered populations (p–r) were the best performers, showing both low bias and much tighter spreads of variance estimates, much like for populations b–e. However with *psill* = 1, for populations j and k, v_8 and v_W showed large underestimation bias, as previously reported by Wolter (2007) for v_8 and by D’Orazio (2003) for v_W , despite being much more precise. For populations l and r with the longest tested autocorrelation *range* (200) under both *psill* levels, v_8 and v_W were both much less biased and much more precise.

One general trend was identified by varying the *range* parameter. At some value of the *range*, between 10 and 50 m for *psill* = 1, and below *range* = 10 for *psill* = 10, all variance estimators shift from below to above the blue line (Fig. 4, h compared to i and m to n), that is, from all underestimating to all, except v_8 and v_W , overestimating the true sampling variance.

Overall, the trends for populations g–r showed similarities to those of populations a–f, with generally lower or much lower bias of v_8 and v_W , always higher precision of v_8 and v_W in more highly clustered populations, and tendency of v_8 and v_W to underestimate, even in some clustered populations (b, c, i, j, k). Between these two best performers, v_W showed greater underestimation bias for some populations (b, c, j, l, q). For all unclustered populations (a, g, h, m), v_8

and v_W underestimated the true sampling variance more than other estimators.

DISCUSSION

True precision of random and systematic survey designs

The square grid systematic design gave variances of the survey mean that were three to five times smaller than the random survey design in populations that clustered by habitat patchiness, aggregating behavior, or both (Table 1, row 5, columns b–f). This implies that a random design would require ~300–500 transects to equal the precision of 100 systematically allocated transects for estimating population density in these clustered populations.

Such large (300–500%) reductions in sampling error variance under a systematic design were not found in the majority of previous studies. One possible reason for this is that this study uses synthetic populations, while many studies used enumerated real world data. Another reason is that we have sought to replicate levels of clustering more commonly observed in marine sampling, while these studies more often used forest or landscape data. However, in Cochran's (1977:223) table of variance ratios, stratified random variances were sometimes two to five times wider than systematic, and all survey design comparison studies we cited have found that systematic designs are more precise than simple random or stratified random. Cochran (1946, see also Iachan 1982) analytically proved that a systematic design is more precise for linear populations whose spatial autocorrelation declines exponentially with separation distance. Examples of systematic designs bringing higher survey precision in spatially autocorrelated populations were reported in a range of applied ecological contexts (e.g., Dunn and Harrison 1993, Ambrosio et al. 2004, Aune-Lundberg and Strand 2014) and in the application of Euclidean distance analysis to animal habitat selection analysis (Benson 2013) and in stereology (Gundersen et al. 1999).

Restricting organisms that are otherwise randomly distributed to habitat patches covering 22% of the survey study area was alone sufficient to yield a high (~300–400%) level of precision improvement with systematic sampling (Table 1, rows 5 and 6, columns b,c). Previous studies comparing random with systematic designs (or comparing the variance of systematic variance estimators) have not investigated the effect of demarcated patchiness in species habitat. Since restriction to favorable habitats is typical of most natural populations, this source of imprecision must be encountered by many environmental and ecological surveys.

In clustered populations, meaningful improvements in survey efficiency permit either lower-cost surveys or more precise information from a given sample

size, affording large increases in statistical power. As the cost constraints on field studies and environmental monitoring increase, meaningful improvements in efficiency favor the choice of a systematic design.

In addition to (1) higher survey precision in clustered populations, two further natural advantages favor a systematic design: (2) systematic designs are more practical to implement. Sample locations positioned on a grid makes systematic designs easier to plan and carry out. Field studies often use systematic designs for this reason. (3) A systematic grid of measurements is superior to random sample locations for use as data input to geostatistical (i.e., contour) mapping.

Spatial mapping inference, modeling the distribution of organisms across the study region, for example by ordinary kriging (Marchant and Lark 2007) or spatial modeling using maximum likelihood (Diggle and Ribeiro 2007), generally give more reliable estimates when survey sample locations are allocated uniformly across the study region, ideally augmented by additional samples at shorter distance scales (Diggle and Ribeiro 2007).

Random sample locations provide an uneven coverage and so a less consistent source of spatial information, and random sampling is not often chosen for mapping applications. In application to abalone fishery management, from surveys using semi-systematic (Byth and Ripley 1980) density measurements, maps of spatial distribution were validated in a fish-down experiment (McGarvey et al. 2008). Contour maps generated from semi-systematic surveys before and after fishing were faithful predictors of where higher abalone densities were confirmed by commercial fishing, and where population density changed most by fishing removals. Geostatistical mapping methods, including kriging (Marchant and Lark 2007, Li and Heap 2008) and spatial statistical modeling (Diggle and Ribeiro 2007) also benefit by information about autocorrelation over shorter separation distances than the distance between systematic samples (Diggle and Lophaven 2006). Subdividing each transect into quadrats and recording organism counts in each quadrat along (within) transects permits shorter (1–100 m) scales of spatial autocorrelation to be measured directly (McGarvey et al. 2010). Other modifications of systematic designs, such as combinations of systematic, stratified, and random (Quenouille 1949) have been proposed.

One practical advantage favors a random design. Harbitz and Pennington (2004) showed that the shortest path among random sample locations traces, on average, a shorter sampling pathway than along a systematic grid. This outcome reflects one objective of systematic sampling, to maintain a wider mean distance between neighboring sample locations, minimizing the effect of spatial autocorrelation, and covering the survey region more uniformly. In

practice, identifying the shortest path among random sample locations can be challenging, particularly for larger sample sizes, there being no general mathematical solution to the traveling salesmen problem. Moreover, in practice, with varied topography and surveys run monthly or yearly over large areas, a shorter path may be a smaller consequence for practical implementation of surveys for researchers in the field than other practical advantages of a systematic design.

Systematic designs probably achieve the greatest precision advantage in pelagic sampling applications. In terrestrial studies, the availability of remote sensing and direct observation means that differences in habitat type can be readily discerned and mapped, permitting reliable stratification of the study region, often greatly improving survey precision. In pelagic sampling using a net, with either vertical or horizontal tows, very high levels of sampling variance are common (Pennington 1983, 1996). Remote sensing can only provide pre-survey knowledge of the surface layer, and the pelagic environment varies over time. High levels of clustering are common in pelagic environments, due to both high patchiness and dense aggregations spread over three dimensions, resulting in differences among sample counts of sometimes two to three orders of magnitude (many samples with zero counts, and a few samples with very high counts of hundreds or thousands of eggs or larvae, or of adults when fish are schooled). Aggregations of copepods over coral reefs were measured in swarm densities of hundreds of thousands per cubic meter (Hamner and Carleton 1979).

Other species known to aggregate include swarming ants (Vandermeer et al. 2008), bees (Fefferman and Starks 2006), and flocking birds (Farley et al. 2008). Other examples include sampling for invasive pests such as locusts (Clark 1969), algal blooms (Ochumba and Kibaara 1989), or more rare species limited to suitable but less well identified habitat (Brown 1984, Ellingsen et al. 2007).

In choosing between a random or systematic allocation of sample locations in environmental monitoring or ecological field work, researchers face a trade-off. In autocorrelated (clustered) populations (identified, in practice, by meaningful numbers of low or zero transect counts and a few big counts), a systematic survey design will nearly always be more precise (Cochran 1946, 1977, D'Orazio 2003). In more highly autocorrelated populations, such as the five clustered populations (b-f) of Fig. 1, the improvement in survey estimate precision with systematic sampling is large. On the other hand, a random design brings more reliably estimated, unbiased confidence intervals using the textbook estimator, v_1 , by meeting the assumption of independence. The trade-off is better (more precise or efficient) estimates with a systematic design in clustered populations, but how (much more) precise remains uncertain.

Testing variance estimators for systematic surveys

To address this uncertainty, following Wolter (1984, 2007), we tested a suite of variance estimation formulas applicable for use with systematic sampling. Some recent studies of systematic variance estimators have limited comparison of the textbook variance estimator to one of several forms of post-stratification estimator (Dunn and Harrison 1993, Aune-Lundberg and Strand 2014) represented in this study by v_2 and v_3 . We examined a wider range of variance estimators than previous studies, evaluating Wolter's eight variance estimators, and five variance estimators designed for use in spatially resolved (not linear) populations, two proposed by D'Orazio (2003), and three covered grid balanced difference estimators proposed here (Appendix).

We did not assess model-based approaches, which show considerable promise (Simmonds and Fryer 1996, Aubry and Debouzie 2000, 2001, Diggle and Ribeiro 2007, Cressie and Wikle 2011, Fewster 2011, Opsomer et al. 2012) but are more difficult to code and implement than the simpler design-based systematic survey variance estimators tested here. Model-based methods depend on a correct choice of model assumptions for the true spatial distribution of organisms inside the survey study region, including its autocorrelation structure (Särndal et al. 1978, Diggle and Ribeiro 2007). This can work very well for studies carried out on restricted sites, but semi-variograms included in Aune-Lundberg and Strand (2014) demonstrate the challenges of fitting models to real-world data for large and variable regions. Modeling methods continue to advance, and Bayesian methods (Cressie and Wikle 2011) or Gaussian random fields across the spatial domain (Thorson et al. 2015) offer further statistical modeling power for surmounting the high and complex variability common in real spatial data.

The other important class of survey methods we did not examine was adaptive sampling (Thompson and Seber 1996, Seber and Salehi 2013). These can be designed for estimating density in clustered populations by placing higher sampling intensity where the species of interest occurs with higher probability as inferred from the survey itself. Generally these must be tailored to each application, and the aim is to find adaptive designs that are unbiased. Asking researchers to implement decisions in the field in real time about where and whether to sample can be challenging, with overestimation a potential risk (McGarvey 2006:93). An alternative method for targeting the higher density sub-areas with higher sampling intensity for which non-bias can be assured by classical statistics is to (pre-) stratify the study region. Strata can be drawn in subsequent iterations of a monitoring survey after a contour map from an initial uniform-grid

systematic survey demarcates areas of high and low density. Systematic sampling within well-drawn strata (Cochran 1977:226–227, Aune-Lundberg and Strand 2014) offers two ways to reduce survey sampling variance.

Among the eight that he tested for systematic samples, Wolter (1984, 2007:332) recommended the two post-stratification variance estimators (v_2 and v_3). Fewster (2011) noted that these have become the most commonly applied (e.g., Millar and Olsen [1995], one by Simmonds and Fryer [1996], Kingsley [2000]) for systematic surveys. However, the simulation results here (Table 1 and Fig. 3) found v_2 and v_3 , after v_1 , to be the most overestimation biased in clustered populations. Similar overestimation bias by v_2 and v_3 was also evident in the 12 populations (i–l, o–r) of Fig. 4, though for these, the differences between v_2 and v_3 and other estimators (besides v_8 and v_W) were less pronounced. D’Orazio’s (2003) v_{STR2} , an adaptation of Wolter’s v_2 , was similarly biased. This poor performance by the post-stratification variance estimators could reflect their having been developed and tested by Wolter for linear, rather than spatially, autocorrelated populations. We did not test the Matérn (1947) two-dimensional versions of local spatial difference estimators used in Scandinavian forest surveys (reviewed by Heikkinen 2006, Aune-Lundberg and Strand 2014).

We found Wolter’s versions of balanced difference estimators (v_4 – v_6) to be less biased than the post-stratification estimators for populations b, c, e, i, l, q, and less clearly so for f, j, o, and p. Covered grid estimators, which we developed as extensions of Wolter’s v_4 – v_6 , improved on v_4 – v_6 in spatially clustered populations b–c, e–f, l, p, q, and more weakly so for i, j, k, o. Covered grid estimators (v_{RCG} , v_{CCG} , v_{TCG}) were the second best performing category of variance estimator.

For clustered populations, covered grid (v_{RCG} , v_{CCG} , v_{TCG}) and other estimators (v_1 – v_6 , v_{STR2}) were considerably more biased, and much less precise, than the best performers among the 13 we tested here, v_8 and v_W . However, for v_W in populations b, c, q, and for both v_8 and v_W in populations i, j, k, underestimation of sampling variance for some clustered populations was evident (Figs. 3 and 4).

These results were effectively reversed in unclustered populations, where all 13 estimators underestimated the true variance (a, h, m), with v_8 and v_W being the most, rather than the least, biased (a, g, h, m). Future work could address the considerable underestimation bias that remains for v_8 and v_W in unclustered populations.

Wolter (2007:332) reported that v_8 has remarkably good properties for the simulated populations with spatial autocorrelation or a linear trend. D’Orazio (2003:293) also concluded that v_W was the appropriate choice for populations with high spatial

correlation. But Wolter and D’Orazio observed underestimation bias of v_8 and v_W respectively, and overall found in favor of post-stratification estimators, notably for lower sample sizes and lower levels of autocorrelation.

Future work: improving on v_8 and v_W

The problem of estimating the variance of the systematic sample mean was first addressed in Scandinavian forestry in the 1920s and 1930s (Heikkinen 2006), and later in North American forestry (Hasel 1938, Osborne 1942, Finney 1948, 1949, Bourdeau 1953). A series of studies by statistical sampling theorists from the 1940s onward, first reviewed by Buckland (1951), extended Cochran’s idea of a superpopulation model to investigate statistical properties of estimators in spatial populations where systematic sampling was more precise (Quenouille 1949, Das 1950, Jowett 1952, Williams 1956, Hannan 1962, Iachan 1982, Bellhouse and Sutradhar 1988). However, this theoretical approach has not yet produced a robust, unbiased variance estimator generally usable with systematic data sets (see e.g., Cochran 1977). Wolter (1984, 2007), D’Orazio (2003), and others proposed and investigated a selection of diverse estimators. The unexpectedly strong performance that we observed for v_8 and its close variant v_W may suggest a pathway forward toward still more reliable variance estimates under systematic sampling.

These two least biased and most precise variance estimators for clustered populations, v_8 and v_W , use formulas that explicitly correct for the observed level of autocorrelation (if positive) between successive systematic samples. In particular, v_8 is written (Wolter 2007:302):

$$v_8 = \begin{cases} (1-f)(s^2/n)[1 + 2/\ln(\hat{\rho}) + 2/(\hat{\rho}^{-1} - 1)], & \text{if } \hat{\rho} > 0 \\ (1-f)(s^2/n), & \text{if } \hat{\rho} \leq 0 \end{cases}$$

where
$$\hat{\rho} = \frac{1}{s^2 \cdot (n-1)} \sum_{i=1}^{n-1} (y_{i+1} - \bar{y}) \cdot (y_i - \bar{y})$$

The autocorrelation $\hat{\rho}$ in v_8 quantifies observed correlation between successive points in the systematic sample grid. In D’Orazio’s v_W (2003), Moran’s index of spatial autocorrelation was substituted for $\hat{\rho}$ in seeking a two-dimensional generalization of v_8 . Future development of better systematic variance estimators in two-dimensional autocorrelated populations should seek to account for the full autocorrelation function vs. separation distance (Quenouille 1949, Matérn 1960, Diggle and Ribeiro 2007) to improve on v_8 and v_W . Relatively low bias by v_8 and v_W for positively autocorrelated (clustered) populations suggests that the $\hat{\rho}$ -dependent correction terms can work well for

organisms clustering by either patchiness of habitat or aggregating behavior, and for the majority of populations clustering by a specified variogram autocorrelation structure. But v_8 and v_W substantially underestimated variance in the unclustered populations.

We conjecture that one important improvement to v_8 and v_W which could reduce this underestimation by v_8 and v_W in unclustered populations would be to extend its reach to permit more direct use of $\hat{\rho}$ estimates that are negative. In Wolter's v_8 , the correction terms are $(2/\ln(\hat{\rho}) + 2/(\hat{\rho}^{-1} - 1))$ applied only for positively autocorrelated (clustered) populations, necessary, in particular, since $\hat{\rho}$ is used as the argument of a log function. In randomly distributed populations, where $\hat{\rho}$ averages around zero, the chance of a sampled negative or positive value is about equal. The $\hat{\rho}$ sign asymmetry in Wolter's v_8 formula specifies that no correction is made to the textbook variance estimator for $\hat{\rho} \leq 0$, which plausibly leads, in part, to the observed underestimation bias of v_8 and v_W in CSR populations. A systematic variance estimator is sought which does not truncate for negative $\hat{\rho}$ and which can incorporate not merely a single scalar $\hat{\rho}$ but the full dependence of autocorrelation on separation distance within the survey study region (Cochran 1946, Quenouille 1949, Diggle and Ribeiro 2007; P. J. Diggle, *personal communication*).

CONCLUSIONS

Systematic survey designs are superior to random designs in clustered populations for three important reasons. They provide more precise survey estimates in populations that cluster by habitat patchiness or aggregating behavior, they are easier to implement, and they produce input data suitable for geostatistical mapping. The better precision of systematic designs for estimating population density should extend to any environmental or ecological quantity that depends on population density such as standing stock biomass, CO₂ absorbed or released, or prey consumed. To address the overestimation of confidence intervals for systematic surveys in clustered populations, among those evaluated, v_8 and v_W were clearly superior for estimating the variance of the survey sample mean, but these were the worst performing in unclustered populations. The statistical search is ongoing for a variance estimator usable with systematic surveys that is unbiased (or meaningfully less biased) in both clustered and random populations.

ACKNOWLEDGMENTS

We thank Peter Diggle for early discussions and recent helpful comments on the draft manuscript. Jason Tanner and Anthony Fowler provided comments on an earlier draft. Michael Pennington provided several useful comments on the draft manuscript. Brian Inouye recommended including additional autocorrelated populations. Sarah Goslee generously provided the R code that she used to generate autocorrelated simulated

spatial point populations, which we adapted to generate the populations of Fig. 4. We also thank two anonymous reviewers for numerous comments that improved the manuscript.

LITERATURE CITED

- Ambrosio, L., L. Iglesias, C. Marin, and J. P. Del Monte. 2004. Evaluation of sampling methods and assessment of the sample size to estimate the weed seedbank in soil, taking into account spatial variability. *Weed Research* 44:224–236.
- Aubry, P., and D. Debouzie. 2000. Geostatistical estimation variance for the spatial mean in two-dimensional systematic sampling. *Ecology* 81:543–553.
- Aubry, P., and D. Debouzie. 2001. Estimation of the mean from a two-dimensional sample: the geostatistical model-based approach. *Ecology* 82:1484–1494.
- Aune-Lundberg, L., and G. H. Strand. 2014. Comparison of variance estimation methods for use with two-dimensional systematic sampling of land use/land cover data. *Environmental Modelling & Software* 61:87–97.
- Babcock, R., and J. Keesing. 1999. Fertilization biology of the abalone *Haliotis laevis*: laboratory and field studies. *Canadian Journal of Fisheries and Aquatic Sciences* 56:1668–1678.
- Baddeley, A., and R. Turner. 2005. Spatstat: an R package for analysing spatial point patterns. *Journal of Statistical Software* 12:1–42.
- Bellhouse, D. R. 1977. Some optimal designs for sampling in two dimensions. *Biometrika* 64:605–611.
- Bellhouse, D. R., and B. C. Sutradhar. 1988. Variance estimation for systematic sampling when autocorrelation is present. *Statistician* 37:327–332.
- Benson, J. F. 2013. Improving rigour and efficiency of use-availability habitat selection analyses with systematic estimation of availability. *Methods in Ecology and Evolution* 4:244–251.
- Bourdeau, P. F. 1953. A test of random versus systematic ecological sampling. *Ecology* 34:499–512.
- Brönmark, C., B. Malmqvist, and C. Otto. 1984. Anti-predator adaptations in a neustonic insect (*Velia caprai*). *Oecologia* 61(2):189–191.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. *American Naturalist* 124:255–279.
- Buckland, W. R. 1951. A review of the literature of systematic sampling. *Journal of the Royal Statistical Society B* 13:208–215.
- Byth, K., and B. Ripley. 1980. On sampling spatial patterns by distance methods. *Biometrics* 36:279–284.
- Clark, D. P. 1969. Night flights of the Australian plague locust, *Chortoicetes terminifera* Walk., in relation to storms. *Australian Journal of Zoology* 17:329–352.
- Cochran, W. G. 1946. Relative accuracy of systematic and stratified random samples for a certain class of populations. *Annals of Mathematical Statistics* 17:164–177.
- Cochran, W. G. 1977. Sampling techniques. Third edition. Wiley, New York, New York, USA.
- Cressie, N., and C. K. Wikle. 2011. Statistics for spatio-temporal data. John Wiley & Sons, Hoboken, New Jersey, USA.
- Das, A. C. 1950. Two-dimensional systematic sampling and the associated stratified and random sampling. *Sankhya* 10:95–108.
- Dew, C. B. 1990. Behavioral ecology of podding red king crab, *Paralithodes camtschatica*. *Canadian Journal of Fisheries and Aquatic Sciences* 47(10):1944–1958.
- Diggle, P., and S. Lophaven. 2006. Bayesian geostatistical design. *Scandinavian Journal of Statistics* 33:53–64.

- Diggle, P. J. 1983. Statistical analysis of spatial point patterns. Academic Press, London, UK.
- Diggle, P. J., and P. J. Ribeiro. 2007. Model-based geostatistics. Springer Science + Business Media, New York, New York, USA.
- D’Orazio, M. 2003. Estimating the variance of the sample mean in two-dimensional systematic sampling. *Journal of Agricultural Biological and Environmental Statistics* 8:280–295.
- Dunn, R., and A. R. Harrison. 1993. Two-dimensional systematic sampling of land use. *Journal of the Royal Statistical Society C* 42:585–601.
- Ehrlén, J., and O. Eriksson. 2000. Dispersal limitation and patch occupancy in forest herbs. *Ecology* 81:1667–1674.
- Ellingsen, K. E., J. E. Hewitt, and S. F. Thrush. 2007. Rare species, habitat diversity and functional redundancy in marine benthos. *Journal of Sea Research* 58:291–301.
- Farley, E. A., K. E. Sieving, and T. A. Contreras. 2008. Characterizing complex mixed-species bird flocks using an objective method for determining species participation. *Journal of Ornithology* 149:451–468.
- Fefferman, N. H., and P. T. Starks. 2006. A modeling approach to swarming in honey bees (*Apis mellifera*). *Insectes Sociaux* 53:37–45.
- Fewster, R. M. 2011. Variance estimation for systematic designs in spatial surveys. *Biometrics* 67:1518–1531.
- Finney, D. J. 1948. Random and systematic sampling in timber surveys. *Forestry* 22(1):64–99.
- Finney, D. J. 1949. An example of periodic variation in forest sampling. *Forestry* 23:96–111.
- Goslee, S. C. 2006. Behavior of vegetation sampling methods in the presence of spatial autocorrelation. *Plant Ecology* 187:203–212.
- Green, R. F., and A. T. Nunez. 1986. Central-place foraging in a patchy environment. *Journal of Theoretical Biology* 123(1):35–43.
- Greig-Smith, P. 1983. Quantitative plant ecology. Volume 9. University of California Press, Berkeley, California, USA.
- Gundersen, H. J. G., E. B. V. Jensen, K. Kiêu, and J. Nielsen. 1999. The efficiency of systematic sampling in stereology—reconsidered. *Journal of Microscopy* 193:199–211.
- Hamner, W. M., and J. H. Carleton. 1979. Copepod swarms: attributes and role in coral reef ecosystems. *Limnology & Oceanography* 24:1–14.
- Hannan, E. J. 1962. Systematic sampling. *Biometrika* 49:281–283.
- Harbitz, A., and M. Pennington. 2004. Comparison of shortest sailing distance through random and regular sampling points. *ICES Journal of Marine Science* 61:140–147.
- Hasel, A. A. 1938. Sampling error in timber surveys. *Journal of Agricultural Research* 57(10):713–736.
- Heikkinen, J. 2006. Chapter 10: assessment of uncertainty in spatially systematic sampling. Pages 155–176 in A. Kangas, and M. Maltamo, editors. *Forest inventory—methodology and applications*. Springer, Dordrecht, Netherlands.
- Iachan, R. 1982. Systematic sampling: a critical review. *International Statistical Review* 50:293–303.
- Jowett, G. H. 1952. The accuracy of systematic sampling from conveyor belts. *Applied Statistics* 1:50–59.
- Kingsley, M. C. S. 2000. Numbers and distribution of beluga whales, *Delphinapterus leucas*, in James Bay, eastern Hudson Bay, and Ungava Bay in Canada during the summer of 1993. *Fishery Bulletin* 98:736–747.
- Legendre, P., M. R. T. Dale, M. J. Fortin, J. Gurevitch, M. Hohn, and D. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* 25:601–615.
- Levine, J. M., and D. J. Murrell. 2003. The community-level consequences of seed dispersal patterns. *Annual Review of Ecology, Evolution, and Systematics* 34:549–574.
- Li, J., and A. D. Heap. 2008. A review of spatial interpolation methods for environmental scientists. *Record* 2008/23. Geoscience Australia, Symonston, Australian Capital Territory, Australia.
- Madow, W. G., and L. H. Madow. 1944. On the theory of systematic sampling, I. *Annals of Mathematical Statistics* 15:1–24.
- Marchant, B. P., and R. M. Lark. 2007. Optimized sample schemes for geostatistical surveys. *Mathematical Geology* 39:113–134.
- Matérn, B. 1947. Methods of estimating the accuracy of line and sample plot surveys. *Meddelanden från Statens Skogsforskningsinstitut* 36:1 [In Swedish with English summary].
- Matérn, B. 1960. Spatial variation: stochastic models and their application to some problems in forest surveys and other sampling investigations. *Statens skogsforskningsinstitut*, Stockholm, Sweden.
- McGarvey, R. 2006. Assessing survey methods for greenlip abalone in South Australia. Final Report to Australian Fisheries Research and Development Corporation Project No. 2001/076. SARDI Aquatic Sciences Publication No. RD01/0147-2. SARDI Research Report Series No. 184.
- McGarvey, R., J. E. Feenstra, S. S. Mayfield, and E. V. Sautter. 2010. A diver survey method to quantify the clustering of sedentary invertebrates by the scale of spatial autocorrelation. *Marine and Freshwater Research* 61:153–162.
- McGarvey, R., S. Mayfield, K. Byth, T. Saunders, R. Chick, B. Foureur, J. E. Feenstra, P. Preece, and A. Jones. 2008. A diver survey design to estimate absolute density, biomass and spatial distribution of abalone. *Canadian Journal of Fisheries and Aquatic Sciences* 65:1931–1944.
- Millar, R. B., and D. Olsen. 1995. Abundance of large toheroa (*Paphies ventricosa* Gray) at Oreti Beach, 1971–90, estimated from two-dimensional systematic samples. *New Zealand Journal of Marine and Freshwater Research* 29:93–99.
- Ochumba, P. B., and D. I. Kibaara. 1989. Observations on blue-green algal blooms in the open waters of Lake Victoria, Kenya. *African Journal of Ecology* 27:23–34.
- Opsomer, J. D., M. Francisco-Fernández, and X. Li. 2012. Model-based non-parametric variance estimation for systematic sampling. *Scandinavian Journal of Statistics* 39:528–542.
- Osborne, J. G. 1942. Sampling errors of systematic and random surveys of cover-type areas. *Journal of the American Statistical Association* 37:256–264.
- Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30:683–691.
- Pebesma, E. J., and C. G. Wesseling. 1998. Gstat: a program for geostatistical modelling, prediction and simulation. *Computers & Geosciences* 24:17–31.
- Pennington, M. 1983. Efficient estimators of abundance, for fish and plankton surveys. *Biometrics* 39:281–286.
- Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. *Fishery Bulletin* 94:498–505.
- Pennington, M., L. M. Burmeister, and V. Hjellvik. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fishery Bulletin* 100:74–80.
- Pennington, M., and J. H. Vølstad. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. *Biometrics* 50:725–732.
- Perry, J. N., A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriri, A. Jakomulska, and S. Citron-Pousty. 2002.

- Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* 25:578–600.
- Quenouille, M. H. 1949. Problems in plane sampling. *Annals of Mathematical Statistics* 20:335–375.
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>
- Särndal, C. E., I. Thomsen, J. M. Hoem, D. V. Lindley, O. Barndorff-Nielsen, and T. Dalenius. 1978. Design-based and model-based inference in survey sampling [with discussion and reply]. *Scandinavian Journal of Statistics* 5:27–52.
- Seber, G. A., and M. M. Salehi. 2013. Adaptive sampling designs: inference for sparse and clustered populations. Springer, Berlin, Germany.
- Seidler, T. G., and J. B. Plotkin. 2006. Seed dispersal and spatial pattern in tropical trees. *PLoS Biology* 4(11):e344.
- Simmonds, E. J., and R. J. Fryer. 1996. Which are better, random or systematic acoustic surveys? A simulation using North Sea herring as an example. *ICES Journal of Marine Science* 53:39–50.
- Stevens, B. G., J. A. Haaga, and W. E. Donaldson. 1994. Aggregative mating of Tanner crabs, *Chionoecetes bairdi*. *Canadian Journal of Fisheries and Aquatic Sciences* 51(6):1273–1280.
- Suzuki, R. O., J. I. Suzuki, and N. Kachi. 2005. Change in spatial distribution patterns of a biennial plant between growth stages and generations in a patchy habitat. *Annals of Botany* 96(6):1009–1017.
- Thompson, S. K., and G. A. F. Seber. 1996. Adaptive sampling. Wiley, New York, New York, USA.
- Thorson, J. T., H. J. Skaug, K. Kristensen, A. O. Shelton, E. J. Ward, J. H. Harms, and J. A. Benante. 2015. The importance of spatial models for estimating the strength of density dependence. *Ecology* 96:1202–1212.
- Vandermeer, J., I. Perfecto, and S. M. Philpott. 2008. Clusters of ant colonies and robust criticality in a tropical agroecosystem. *Nature* 451:457–460.
- Williams, R. M. 1956. The variance of the mean of systematic samples. *Biometrika* 43:137–148.
- Wolter, K. M. 1984. An investigation of some estimators of variance for systematic sampling. *Journal of the American Statistical Association* 79:781–790.
- Wolter, K. M. 2007. Introduction to variance estimation. Second edition. Springer Science + Business Media, New York, New York, USA.
- Yates, F. 1948. Systematic sampling. *Philosophical Transactions of the Royal Society A* 241(834):345–377.
- Yates, F. 1960. Sampling methods for censuses and surveys. Third edition. Charles Griffin & Company, London, UK.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1890/14-1973.1/supinfo>