Landscape Ecology **18:** 741–757, 2003.

Research article

The influence of sampling scheme and interpolation method on the power to detect spatial effects of forest birds in Ontario (Canada)

Robert S. Rempel* and Robert S. Kushneriuk
Ontario Ministry of Natural Resources, Centre for Northern Forest Ecosystem Research, Lakehead University, Thunder Bay, Ontario, Canada, P7B 5E1; **Author for correspondence (e-mail: rob.rempel@mnr.gov.on.ca)*

Received 24 October 2002; accepted in revised form 17 April 2003

Key words: Habitat associations, Interpolation, Kriging, Landscape ecology, Power analysis, Spatial accuracy assessment, Spatial ecology, Songbird population abundance

Abstract

Spatial ecology is becoming an increasingly important component of resource management, and the general monitoring of how human activities affect the distribution and abundance of wildlife. Yet most work on the reliability of sampling strategies is based on a non-spatial analysis of variance paradigm, and little work has been done assessing the power of alternative spatial methods for creating reliable maps of animal abundance. Such a map forms a critical response variable for multiple scale studies relating landscape structure to biotic function. The power to reconstruct patterns of distribution and abundance is influenced by sample placement strategy and density, the nature of spatial auto-correlation among points, and by the technique used to extrapolate points into an animal abundance map. Faced with uncertainty concerning the influence of these factors, we chose to first synthesize a model reference system of known properties and then evaluate the relative performance of alternative sampling and mapping procedures using it. We used published habitat associations of tree nesting boreal neo-tropical birds, a classified habitat map from the Manitou Lakes area of northwestern Ontario, and point count means and variances determined from field studies in boreal Canada to create 4 simulated models of avian abundance to function as reference maps. Four point sampling strategies were evaluated by 4 spatial mapping methods. We found mixed-cluster sampling to be an effective point sampling strategy, particularly when high habitat fragmentation was avoided by restricting samples to habitat patches > 10 ha in size. We also found that of the 4 mapping methods, only stratified ordinary point kriging (OPK) was able to generate maps that reproduced an embedded landscape-scale spatial effect that reduced nesting bird abundance in areas of higher forest age-class fragmentation. Global OPK was effective only for detecting broader, regional-scale differences.

Introduction

Spatial ecology is becoming an increasingly important component of resource management, and the general monitoring of how human activities affect the distribution and abundance of wildlife. Most work on the reliability of sampling strategies is based on a non-spatial analysis of variance paradigm, and a considerable body of literature has developed to estimate the optimal design and reliability of non-spatial wildlife monitoring data (e.g., Hayes and Steidl 1997; Steidl et al. 1997). Such studies have included the simulation of environmental impacts, and analysis of statistical power using Monte Carlo techniques (e.g., Benedetti-Cecchi 2001), and have also examined sampling strategies specific to forest monitoring (Foster 2001) and forest songbird analysis (Carlson 2001; Thompson et al. 2002). For traditional power analysis, the general aim is to determine the appropriate sampling strategy and sample size to reliably answer questions, given an estimated variance and arbitrarily determined effects size. Under the analysis of variance approach, spatial auto-correlation is an unwanted complication that violates the independence of individual observations. Yet in landscape ecology, this violation almost always occurs because vegetative patterns, and associated habitat patterns, are driven by contagious and non-random patterns of environmental factors and natural disturbance (Foster et al. 1998; Dale and Fortin 2002; Fortin and Payette 2002; Ryan 2002). Indeed, the creation of a contiguous surface-response or contour map of animal distribution is best accomplished when the degree and form of spatial auto-correlation among sample points is known.

Hence, in the field of spatial ecology, the challenge is to determine the appropriate sampling scheme (where scheme is defined as the combination of sampling strategy and sample point density) and mapping approach required to estimate spatial auto-correlation, and to create a map that reliably reproduces an initial spatially defined variance pattern (e.g., Fortin et al. 1989; Rossi et al. 1992; Legendre 1993; Bellehumeur and Legendre 1998; Gunnarsson et al. 1998; Dessard 1999; Roy and Tomar 2000; van Groenigen 2000; Hunsaker et al. 2001; Lin and Rouhani 2001; Lichstein et al. 2002; Venier et al. 2002). The emerging disciplines of spatial accuracy assessment and the statistical analysis of spatial data reflect the general importance of this issue (e.g., Lowell and Jaton 1999; Liebhold and Gurevitch 2002). A spatial abundance map forms an important exploratory tool for monitoring population trends in changing landscapes, and makes a useful response variable for multiple scale studies relating landscape structure to biotic function. This is especially important when a compartmental approach to multiple scale habitat modelling is used, where a hierarchy of window sizes is used to estimate landscape pattern indices and associated average species abundance within windows (e.g. Potvin et al. 2001). The sampling scheme itself presents several challenges as landscape patterns can be structured hierarchically (Hall et al. 1988; Kotliar and Wiens 1990; Royle and Berliner 1999; Venier et al. 1999; Boone and Krohn 2000; Elkie and Rempel 2001), and spatial variance across spatial scales may be found in complex, nested patterns (Legendre and Fortin 1989; Bellehumeur and Legendre 1998; Meisel and Turner 1998).

For spatial mapping, two critical uncertainties are: (1) the best sampling scheme to estimate spatial auto-correlation, and (2) the best interpolation technique to reliably map the patterns of distribution and abundance. These factors interact, so they should be addressed in a factorial manner. Our approach was to first create a model system with a spatially defined variance surface (simulated map of bird abundance), and then evaluate the relative performance of alternative sampling and mapping procedures. The modeling approach described here provides an effective means of exploring and contrasting the strengths of alternative strategies. This becomes especially important for more complex sampling and mapping problems, where the interaction of errors complicates their theoretical modeling.

The objectives of the research are to: (1) develop and demonstrate a simulation modeling approach for comparing and assessing the performance of alternative sampling and mapping systems, where the impetus is to create a spatially accurate map depicting the pattern of abundance, and (2) compare the performance of alternative mapping procedures for detecting spatially explicit, landscape-level effects on bird distribution and abundance. The reference point for comparison is a simulated spatial distribution of bird abundance based on habitat specific means and variances from the Manitou Lakes area of northwestern Ontario.

Methods

This study generated 4 simulated reference maps of bird density, differing only in their variance-mean ratio, to evaluate the sampling and mapping systems. Point sample datasets were obtained from the maps by applying 4 sampling strategies in combination with an additional spatial restriction (based on minimal patch area or edge distance). Following this 4 mapping approaches were used to create grids of bird density. The performance of the mapping approaches was evaluated by using 2 formal performance indices, semi-variance model validation and point kriging cross-validation, as well as regression-based measures of map correspondence. Technical details of these methods follow. To demonstrate properties of the simulation modeling system 8 case studies were constructed, using selected combinations of these methodologies. The description and motivation of these case studies follows as well.

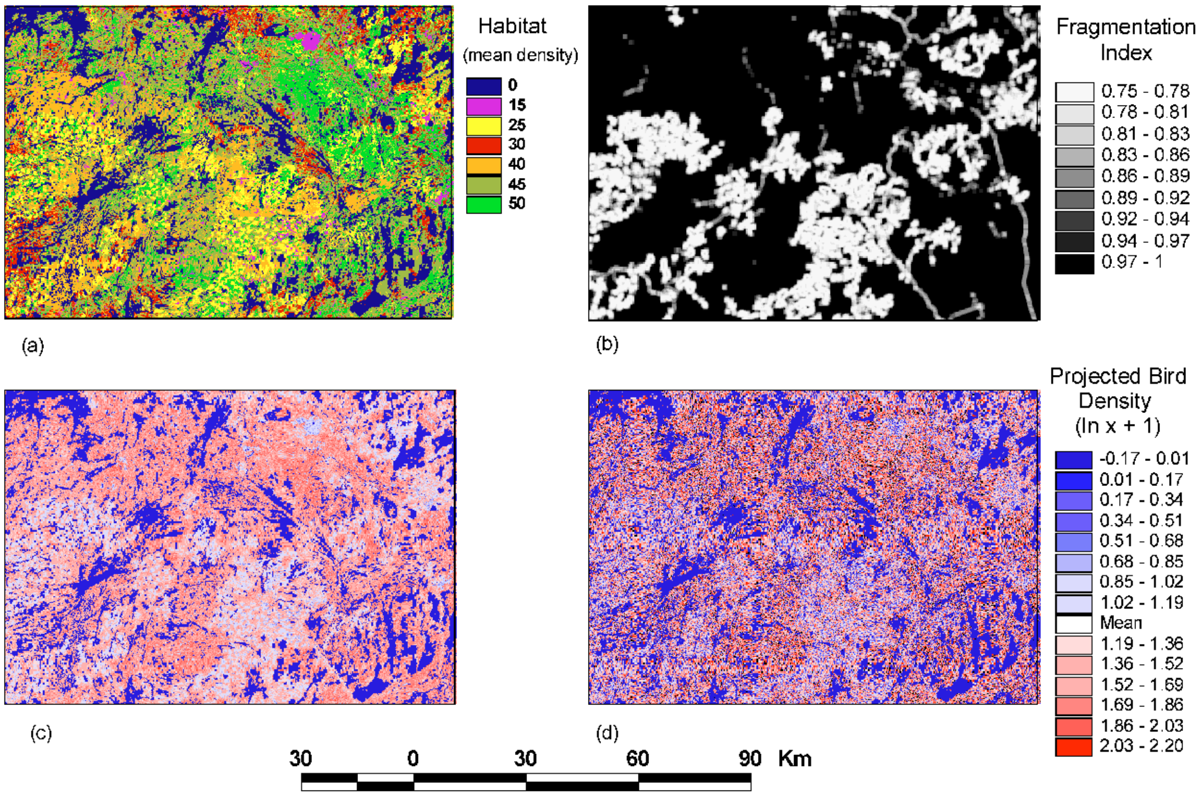

Figure 1. Sequence to create reference bird count simulations. (a) Landsat image classified into 7 habitat classes. For simulations, h = habitat class / 10; (b) fragmentation map, where values (f) between 0.75 − 1.0 represent the variance of young and old age class edges within a 500-m moving average window; (c) reference map LV2, where values represent log-transformed number of birds/ha (i.e., 100 m raster cell), with mean = 1.53 and s^2 = 0.074; (d) reference map HV2, where mean = 1.41 and s^2 = 0.379. Colors for maps (c) and (d) range from cyan (lowest density) to dark red (highest density), with each color separated by 0.5 standard deviation.

Simulated reference maps

We generated simulated reference maps of breeding bird abundance by modeling abundance based on habitat (landcover) associations. Both mean abundance and variance were defined as functions of landcover type (h), and 4 alternative models were created. These consisted of two low variance models (LV1, LV2), where variance for the random model was 0.5 and 1 times the mean abundance for the landcover type, and two high variance models (HV1, HV2) with coefficients of 2 and 4. The 7 landcover types, classified by satellite imagery (Figure 1a), were assigned variable densities of 0 to 5 counts/ha in a manner consistent with known habitat associations of tree nesting neo-tropical birds.

This initial model can be fully specified without any knowledge of the spatial arrangement of habitat patches, and thus over simplifies the complex interaction between a species and its habitat. For example, high interspersion between young and older forest age-classes (fragmentation) is known to reduce bird densities (e.g., Andren 1994; Austen et al. 2001; Boulinier et al. 2001). To investigate this landscape-level effect, a fragmentation thematic layer (*f*) was created by running a 500 m neighborhood analysis of a binary map containing only young ($<$ 20 years) and old forest-age classes (Figure 1b). The observed values ranged between 0–0.25 within the Manitou boreal landscape, and these values were used to depress the mean abundance value as a direct proportion.

The following random model was used to generate a map at a pixel resolution of 100 meters:

$$m_{xy;s} = r_{xy}(h_{xy}, \sqrt{(v_{xy;s})}) \cdot (1.0 - f_{xy})$$

where xy are the map column and row coordinates, s represents the 4 low and high variance scenarios, r_{xy} is a random normal deviate with a landcover specific mean h_{xy} and standard deviation of $\sqrt{(v_{xy;s})}$, f_{xy} is the cell value of the fragmentation thematic layer, and $m_{xy;s}$ is the resulting cell value for map s. Significant digits on internal calculations were maintained before final truncation to integer values; maps values < 0 were set to 0.

$v_{xy;s}$ for maps $s = 1 \ldots 4$ is defined as:

$$v_{xy;1} = h_{xy} * 0.5 \quad \text{(LV1)}$$

$$v_{xy;2} = h_{xy} * 1 \quad \text{(LV2)}$$

$$v_{xy;3} = h_{xy} * 2 \quad \text{(HV1)}$$

$$v_{xy;4} = h_{xy} * 4 \quad \text{(HV2)}$$

The range of means and variances were established by reviewing values for several species from a review of breeding bird point-count studies in the boreal forest (Hobson and Schieck 1999). Bird counts at a listening station cannot be less than zero, so map values < 0 were set to 0. This resulted in a log-normal distribution of the count values; to remedy this situation $m_{xy;s}$ values were transformed to re-establish a normal distribution using the negative binomial transformation (Elliott 1977):

$$m'_{xy;s} = \log_e(m_{xy;s} + 1.0)$$

Although the truncation depressed the initial within class variation, we reduced the effect of the truncation by creating the range of variance scenarios, $m1$ – $m4$. The transformation is also efficient at removing the variance-mean heteroscedasticity defined during the model formulation. Count data typically requires a variance stabilizing transformation, and although the repulsed distribution of territorial sites suggests that a transformation of lesser severity may be sufficient, studies to date indicate that the applied transformation is appropriate (e.g., Hobson et al. 2000).

Sampling strategies

We investigated 4 sampling strategies (note that open water was never sampled for any strategy). A computer program was written by one of us (RK) to generate the point sample dataset using any of the various options outlined below:

(i) Random: Points were selected without respect to landcover type.

(ii) Systematic: Points were selected as the center of a 2500 ha hexagonal grid, without respect to landcover type.

(iii) Mixed-cluster: This strategy combines elements of the staggered systematic, clustered, and random strategies. Points were selected in clusters, with one cluster in each element of a 2500 ha hexagonal grid. The starting point for each cluster was randomly selected within a 2000 m radius of the hexagon centroid, and subsequent points were randomly taken within 1000 m of this start location (Figure 2). Points were selected without respect to area of eligible landcover types. This formal definition originates here, and was motivated by the desire to create an evenly spaced distribution, but at the same time avoid placement of points in non-eligible classes such as water.

(iv) Modified stratified-random: Points were positioned with respect to the proportions of landcover types at the neighborhood level, rather than the global level. We used a compartmentalized approach to stratification, where we first overlaid a hexagonal grid, and then determined the area of each landcover type (strata) within each hexagon grid cell (compartment). The stratification grid was a 2500 ha hexagon grid, with cell centers separated by 5373 m, which we nominally refer to as 5000 m. The number of points within each grid cell was initially set to 5 points, but was adjusted by the proportion of non-water area; water being the only non-eligible habitat type. These points were then assigned in proportion to each habitat type (i.e., if 70% of the landcover was mixedwood, then 70% of the points would be placed within the mixedwood type). Assigned points were placed randomly in the designated habitat. Again, this formal definition of the strategy originates here.

Two other sampling restrictions were also investigated: a restriction on minimum patch size, and a restriction on sampling close to stand boundaries (e.g., patches a minimum of 10 ha, and points at least 100 m from a stand edge). These restrictions were applied to each of the sampling strategies, and for each sampling strategy, a range of synthesized sample point densities was investigated.

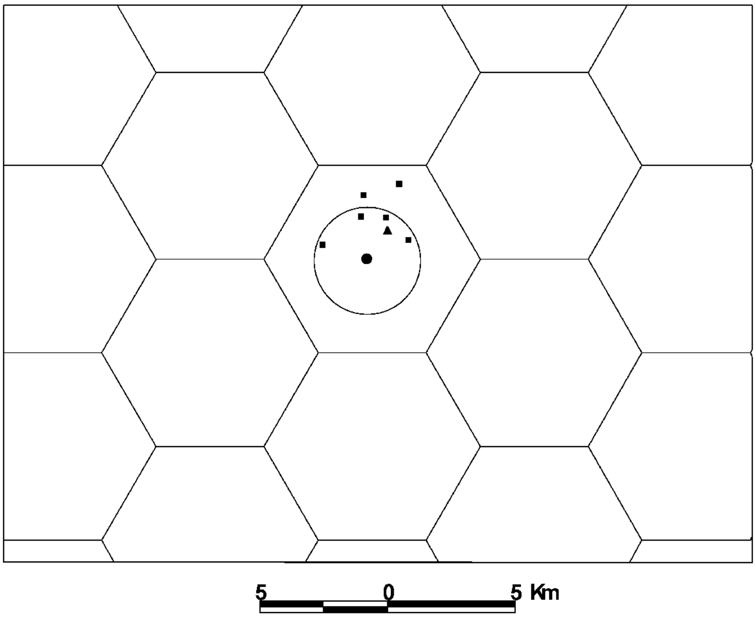

Figure 2. Mixed-cluster sampling. A random point (triangle) is selected within 2,000 meters of the hexagon center (round), then 5 sample points (square), are randomly selected within 1,000 m of the triangle.

Mapping approaches

The effect of mapping procedure on landscape-level density estimates was determined by comparing 4 different mapping approaches: (i) the non-spatial assignment of estimated mean abundance values to each habitat type occurring in the Landsat-based landcover map, (ii) spatial interpolation through ordinary point kriging (OPK), (iii) spatial interpolation through OPK with habitat as a covariate, and (iv) spatial interpolation through stratified OPK with habitat as the stratification criteria. We follow Wallerman et al. (2002) in terming OPK "global OPK" when all points are used. All mapping approaches used synthesized points generated by the mixed-cluster strategy with a 5000 m spacing of 5-point cluster samples.

Stratified OPK is a 4-step procedure: (i) sample points are overlaid on the Landsat-based landcover map, and the habitat type underlying the point is determined; (ii) in turn, an interpolated surface is created from the subset of points corresponding to each landcover type; (iii) conditional modeling is used in the GIS to determine, for each cell, what habitat type it belongs to, and subsequently the appropriate interpolated surface to look up the abundance value, and (iv) those grid cells which fall beyond the defined range of spatial autocorrelation from observed data points (i.e., no-data holes within the kriged map) are assigned the expected (mean) value for that habitat type. Note that OPK with covariates modifies the expected (interpolated) value based on the relationship between habitat type and sampled bird abundance, whereas stratified kriging uses OPK to create separate surfaces of bird abundance for each habitat type.

Performance indices

Semi-variance model validation, point kriging cross-validation, and regression models were used to determine how well the sampling strategies modeled spatially explicit patterns of distribution and abundance.

Semi-variance model validation provides an index of how well spatial autocorrelation is modeled, which in part determines our ability to recreate the spatial patterns of abundance from the point sample data. With only a few, sparsely distributed points, spatial auto-correlation among data points will be modeled less well than if point density is relatively higher (where point density is number of sample points per unit area). Spatial-autocorrelation was modeled by fitting various standard model forms to the semi-variogram, (Goovaerts 1997), and then choosing the model form with the best fit.

Point kriging cross-validation was used to assess the performance of the kriging interpolation. This is a jack-knife cross-validation procedure whereby a sample point is removed, the map re-interpolated at the location of the removed sample point, and the cell value at the removed sample point estimated from the interpolated map (Davis 1987). When repeated for each sample point, a regression of observed versus calculated values can be created. For a well modeled spatial relationship, the slope of the regression and the proportion of accounted variance (r^2) will both approach 1. We chose r^2 as a performance measure for its simplicity and ease of interpretation. Although there are other performance statistics available, none is without limitation (Goovaerts 1997). The performance measures were applied here as an exploratory comparative tool among the various sampling schemes, and not as a rigorous test for “best model”.

Regression models, both simple and multiple, were used to establish the degree of correspondence between the reference and interpolated maps, on a cell-by-cell basis.

Case studies

Eight case studies were used to demonstrate the properties of the simulation system, and results of the 8 cases are presented in Results sections 3.1 − 3.8, respectively.

Case 1. presents the underlying characteristics and spatial structure of the LV1, LV2, HV1 and HV2 simulated reference maps. Sill values and the proportion of structured variance were examined after application of OPK from ten runs of 2880 randomly placed points. Subsequent analyses used only the LV2 and HV2 reference maps.

Case 2. evaluates the effect of edge proximity and patch size on mapping performance by applying the mixed-cluster strategy using a 5000 m spacing, a 2000 m primary radius, 1000 m secondary radius and 5 points per cluster. Samples were eliminated from patch sizes less than 10 ha or edge proximities of less than 100 m. Based on these results all subsequent analyses applied both spatial restrictions.

Case 3. evaluates each of the 4 sampling strategies on ca. 2000 points drawn from the LV2 reference map. Following OPK the semi-variance and cross validation statistics were used to evaluate the performance. Our variances did not follow a normal distribution (Zar 1984), so hypothesis testing of regression r^2 values was generally conducted using a non-parametric test (Kruskal-Wallis k-independent samples). In several cases the principal question was related to main effects versus interactions, so fixed-effects analysis of variance was used to explore interactions, recognizing that the p-values may be biased. The p-values were used to evaluate trends in performance, not to conduct inferential hypothesis testing.

Case 4. evaluates the effects of sample density on mapping performance. We evaluated mixed cluster sampling on the LV2 reference map across five sampling densities (1,000 − 2,000 m spacing of 5 points clusters). OPK was used for interpolation, and the semi-variance and cross-validation techniques were the performance measures.

Case 5. evaluates the power of the 4 mapping approaches (see above) to effectively detect a landscape-level spatial effect – the depression of population abundance due to fragmentation of the landscape. This case is essentially an example application of spatial modeling, where two landscapes, one fragmented, the other not, were compared to the reference map in terms of their estimated mean population density. Recall that the simulated bird abundances were modified by a fragmentation layer, where densities were decreased from 0 − 25%, based on the degree of fragmentation. Performance was based on the deviation between the true and estimated weighted means within the fragmented versus un-fragmented landscapes. Analysis was restricted to deciduous, mixedwood, and conifer mature forest types.

Case 6. evaluates the influence of assessment scale versus mapping approach. Here we characterized the difference between alternative mapping approaches (e.g., global OPK and stratified OPK) in representing fine scale versus broad scale patterns in the landscape. To do this, we overlaid 9 hexagonal grid layers, with grid cells ranging from 10 − 33670 ha, on both the reference and the interpolated maps, and then calculated the weighted average of abundance within each hexagonal grid cell (Figure 3). This procedure was repeated 6 times by offsetting the grid in such a manner as to minimize the overlap held in common over all 6 trials, and then re-calculating the regression parameters and their mean values. A computer program to automate this procedure was written by one of us (RK). Under this analysis approach, data smoothing occurs across a range of spatial scales, so it is possible to evaluate the spatial scale at which the map performs most effectively, and scales at which different mapping approaches converge or diverge with respect to performance indices.

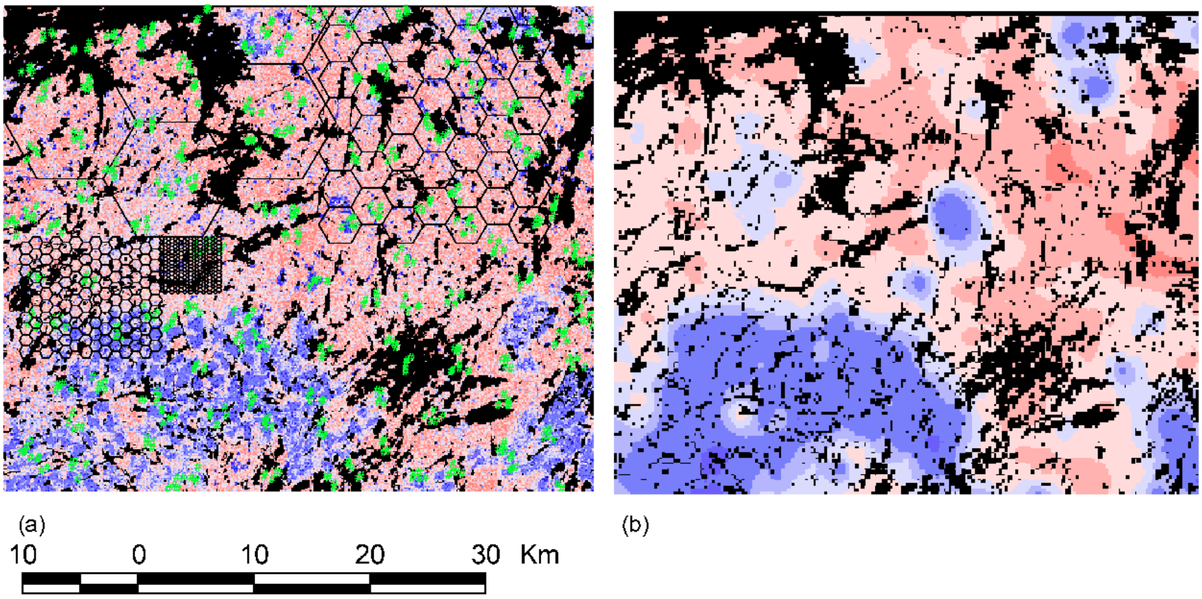

Figure 3. Comparison of reference LV2 basemap (a) with interpolated map (b). Colors as in Figure 1c. Hexagon overlays of 10000, 1000, 100, and 10 ha grids illustrate the major scales for the grid analyses, where analyses are based on regressions of the weighted average of density within each grid cell for the base and interpolated maps. Mixed-cluster sampling is illustrated, where clusters are nominally 5000 m apart, with all 5 points of a cluster within 1000 m of the cluster start point (see Figure 2).

The overall correspondence between the reference maps and interpolated maps was evaluated by determining the weighted mean within each hexagon, and through simple regression relating the two values:

$$r = m * k_i + b$$

where r is the weighted log-transformed mean abundance (loge (x+1)) of birds within the reference map, k is the same for the kriged map, i is the hexagon cell number, m is the slope, and b the intercept. A well-modeled relationship will approach a slope of 1 (low bias), and r^2 of 1 (high precision). Once again comparisons were made among the sampling strategies.

Case 7. evaluates the effects of sample point density on the power of stratified OPK to detect a landscape-level spatial effect. The evaluation is based on a multiple linear regression analysis of accuracy versus inter-cluster spacing, number of points within a cluster, and total number of points. The accuracy index was defined as the deviation from the estimated mean population density (loge (x+1)) of birds found in the reference fragmented landscape versus the interpolated fragmented landscape. The deviation, or error in the estimate of the mean is:

$$\Delta = \sqrt{(x - y)^2}$$

where Δ is the distance between weighted means, and x and y are the weighted mean bird densities within the landscapes for the reference and interpolated maps, respectively. Comparisons were made for the both the high- and low-variance simulated abundance maps HV2 and LV2.

Sampling densities were compared using mixed-cluster sampling, with clusters ranging from 2000 − 10,000 m spacing, and with within cluster point counts of 3 and 5. We were unable to find kriging solutions for 3-point clusters at the 10,000 m spacing level, so this combination was dropped from the analysis.

Case 8 evaluates the practicality and logistics of implementing principles of the preferred sampling scheme. The best computer-generated sampling strategy was applied in an actual field study.

Table 1. Summary statistics within individual habitat types for base reference map, LV2, following loge (x+1) transformation.

Habitat Type (h)	Initial target mean[1].	Final map mean[2]	Var	CoV	Area (ha)
Water (0)	0	0	0	221211	
Wetlands (1.5)	0.92	0.866	0.028	3.18	22872
Recent Disturbance (2.5)	1.26	1.064	0.019	1.76	87909
Deciduous (3.0)	1.39	1.338	0.023	1.73	71020
Mixed Deciduous (4.0)	1.61	1.538	0.023	1.48	174565
Mixed Conifer (4.5)	1.70	1.658	0.020	1.22	280328
Conifer (5.0)	1.79	1.721	0.021	1.22	141327

[1]Initial target mean, before application of landscape-level fragmentation effects (loge ((h) + 1)).
[2]Map mean values after application of fragmentation effects.

Table 2. Summary of variance estimates (across all habitat types) from reference maps, sample points[1], and semi-variance analysis[1], following loge (x+1) transformation.

	Mean		Variance		Semi-Variance Analysis		
	Map	Sample	Map	Sample	Sill	Nugget	Prop2
LV1 exp	1.53	1.52	0.062	0.071	0.071	0.011	0.849
LV2 exp	1.52	1.51	0.074	0.084	0.083	0.025	0.700
HV1 linear	1.50	1.49	0.133	0.144	0.159	0.114	0.284
HV2 linear	1.41	1.40	0.379	0.388	0.402	0.359	0.100

[1]Values based on means of 10 sampling simulations of 2280 points each.
[2]Proportion of variance attributable to spatial structure

Results

Spatial structure of simulated reference maps

As expected, the application of fragmentation effects resulted in a consistently lower mean ($\log_e$ (x+1) transformed) landcover values than the starting values, as illustrated by the summary statistics for map LV2 (Table 1). Sill values of the selected semi-variance models (where sill = structured + unstructured variance) differed little from overall point-sample or reference-map variance, so the semi-variance models were performing well with respect to modeling overall variance. There was no evidence of anisotropic effects for any of the maps; so all kriging analyses were based on isotropic models.

For the low variance reference maps (LV1 and LV2), the proportion of structured variance was 0.85 and 0.70, respectively (Table 2). The relatively high degree of structural variance detected in these maps is directly related to the spatial organization of the landcover types. Landcover types (open water, mature forest, immature forest, etc.) are not organized randomly in space, but are the consequence of highly structured, contagious events on the landscape (e.g., wildfire and geo-fluvial processes). Hence much of the detected spatial pattern in the simulated bird abundance is a consequence of the non-random association of landcover types (as mapped by classified Landsat TM imagery) that were used for generating the reference maps for the study area. As map variance increased, the ability to detect structural variance decreased. The proportion of structural variance detected dropped to 0.284 and 0.1 for maps HV1 and HV2, respectively. High inter-site (i.e., inter-pixel) variability (see Figure 1d) reduces the ability to model underlying structural variance.

The semi-variance in the two low-variance maps was modeled with an exponential model, whereas the higher variance maps (HV1 and HV2) followed a linear model (Table 2). Goodness-of-fit of the semi-variance model was indicated by r^2, and the two low variance maps performed substantially better than the high variance maps ($\chi^2 = 29.3$; df = 1; $p < 0.001$) (Table 3). The effective range of estimated spatial auto-correlation was about 6 times greater with the low variance maps, and at 60,000 m, is approximately equal to the extent of the study area. It is noteworthy that a variety of non-linear models (e.g., spherical, exponential, and Gaussian) performed almost equally well for the low-variance maps, but we found that only the linear model provided a reasonable estimate

Table 3. Performance of semi-variance model1 (model fit) and point kriging1 (cross-validation) for reference maps with increasing among-cell variance.

Reference map	Map s^2	Model fit (r^2)	Cross-validation r^2
LV1	0.062	0.940	0.727
LV2	0.074	0.936	0.563
HV1	0.133	0.529	0.181
HV2	0.379	0.452	0.038

[1]Values based on means of 10 sampling simulations of 2280 points each.

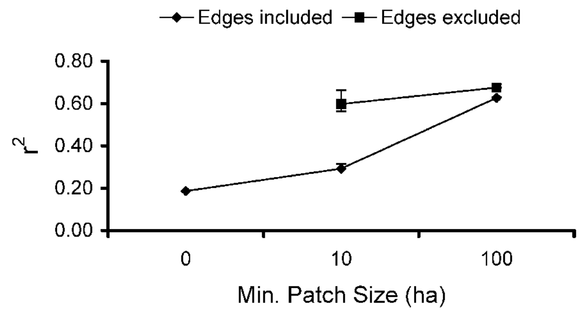

Figure 4. Effect of excluding edges, and constraining minimum patch size, on mapping (cross-validation r^2) performance. Points were selected using mixed-cluster sampling (5000 m spacing between clusters, 1000 m radius of points within a cluster). Error bars are 2nd and 3rd quartiles.

of sill variance for the high-variance maps. The linear model assumes a constant rate of change in spatial auto-correlation across the entire range of sample points, and is reflective of a less spatially structured distribution.

Using the exhaustive point samples, we calculated kriging cross-validation r^2 for the 4 maps to determine how among-cell variance in the reference maps would be expected to affect kriging performance. Cross-validation of sample points was related to map variance, with r^2 varying from 0.727 (LV1) to 0.038 (HV2) (Table 3). Kriging performance was significantly higher for the 2 low variance maps than the 2 high variance maps ($\chi^2 = 29.7$; df = 1; $p < 0.001$). Further analyses were restricted to reference maps LV2 and HV2.

Edge proximity and patch size effects

Both placement of samples with respect to distance from edge, and minimum patch size affected sample performance. Eliminating samples < 100 m from habitat edge improved mapping performance (Figure 4), as did increasing minimum patch size. The edge sampling restriction, however, also effectively resulted in a minimum patch size restriction because edges are defined by 1 ha pixels. Thus patches with sufficient internal core area to permit sample placement away from an edge are all > 9 ha. At the 10 ha min. patch size the "edges excluded" strategy has better performance than the "edges included" strategy, but at the 100 ha min. patch size there is no difference between strategies. Thus the edge restriction has an effect within smaller, irregularly shaped, patches but is insignificant for larger patch sizes where few points would randomly fall within the proximity threshold.

Sampling Strategy and Mapping Performance

Mixed-cluster sampling outperformed random, systematic, and stratified-random sampling (Table 4) in that this sampling strategy had a higher cross-validation index (k-w $\chi^2 = 35.5$; df = 3; $p < 0.001$). Spatial auto-correlation (structural variance) was modeled similarly for all 4 methods (Table 4), but systematic sampling performed slightly poorer than mixed-cluster, random, and modified stratified-random (k-w $\chi^2 = 11.6$; df =3; $p < 0.009$). The sample point mean for mixed-cluster sampling (1.51), however, was less biased and much closer to the reference map mean (1.52) than for the other 3 sampling strategies (k-w $\chi^2 = 20.2$; df = 3; $p < 0.001$). The lower slope values for the systematic, random, and modified stratified-random sampling strategies (Table 4) further reflect this.

Sample point density and mapping performance

Sample point density had a dramatic effect on mapping performance. Sample performance decreased almost linearly over the range 1000 − 5000 m (Figure 5), and variance increased by orders of magnitude for the 10,000 and 20,000 m spacing.

Mapping approach and power to detect spatial effects

The 4 different mapping approaches used to compare bird density between a fragmented and un-fragmented landscape (Figure 6a-b) resulted in substantially different estimates of population density. The spatial only approach exaggerated the differences between

Table 4. Effect of sampling strategy on performance indices[1].

	Spacing (m)	N	Prop of structural variance	SV r^2	Kriging r^2	Slope	Sample[2] Mean
Mixed-cluster	5325	2025	0.732	0.942	0.590	1.006	1.511
Systematic	720	2009	0.698	0.983	0.408	0.959	1.492
Random	Variable	2020	0.756	0.980	0.460	0.957	1.490
Stratified − Random	Variable	2121	0.727	0.942	0.362	0.916	1.495

[1]Values based on means of 10 sampling simulations.
[2]Reference map mean = 1.52.

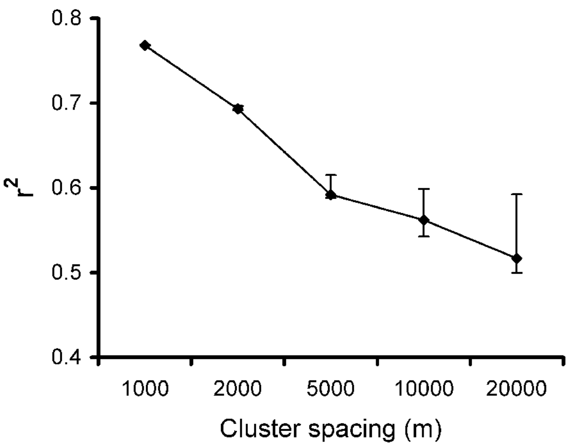

Figure 5. Effect of sample spacing (intensity) on mapping performance. Performance indices (cross-validation r^2) are for 5-point cluster sampling of LV2; sample size varied between 60 − 16,785 for sample spacing over the range 20,000 − 1,000 m. Values are averages for 5 simulated point-sample selection runs.

landscapes by underestimating the true values, and densities within the 2 landscapes varied considerably from the reference means (Table 5). The habitat only approach had insufficient sensitivity to detect spatial effects; so estimated density differed only marginally between landscapes for the mature forest habitat type (Table 5), and was overestimated for the fragmented landscape. Co-kriging improved the estimates of mean population density slightly over global OPK, but still created a negatively biased estimate for the fragmented landscape. Stratified OPK performed markedly better, and accurately detected the difference in mean values for the fragmented and un-fragmented landscapes (Table 5).

Assessment scale versus mapping performance

We found that for high variance and low variance reference maps, stratified OPK outperformed global kriging, and that the difference in performance was greatest at finer spatial scales (Figure 7). For broader scale comparisons ($>$ 360 ha), performance for the 2 techniques began to converge. For the high variance HV2 simulation, the performance response was not linear across scales, and for both techniques, performance peaked at the 30 − 70 ha scale.

Sample point density and power to detect spatial effects

Partial correlation coefficients for the multiple linear regression analysis of the accuracy index Δ versus inter-cluster spacing, number of points within a cluster, and total number of points, reveal that inter-cluster spacing is the most important variable for predicting accuracy within this study. The plot of inter-cluster spacing versus Δ reveals that error increases almost linearly as inter-cluster distance increases, and is further increased if number of points within a cluster is reduced from 5 to 3 (Figure 8). This pattern is similar to the response of sample cross-validation conducted for global OPK (Figure 5), and suggests the decrease in performance for stratified OPK is a result of the composite errors of the separately kriged layers that comprise the stratified OPK result. Similarly, the plot of error versus total sample size reveals that both error in individual runs, and overall variance among runs, decreases as sample size increases (Figure 9).

Field test of sampling strategy logistics

To test the feasibility of the mixed-cluster sampling strategy we attempted to implement the computer-based strategy in a study of breeding boreal forest birds in northwestern Ontario. The logistical difficulties imposed by forest access prevented us from strictly following the computer generated sampling design. Instead, we used the general principles generated from the simulation studies to approximate the sampling strategy. We found a close approximation of the computer generated mixed-cluster sampling can

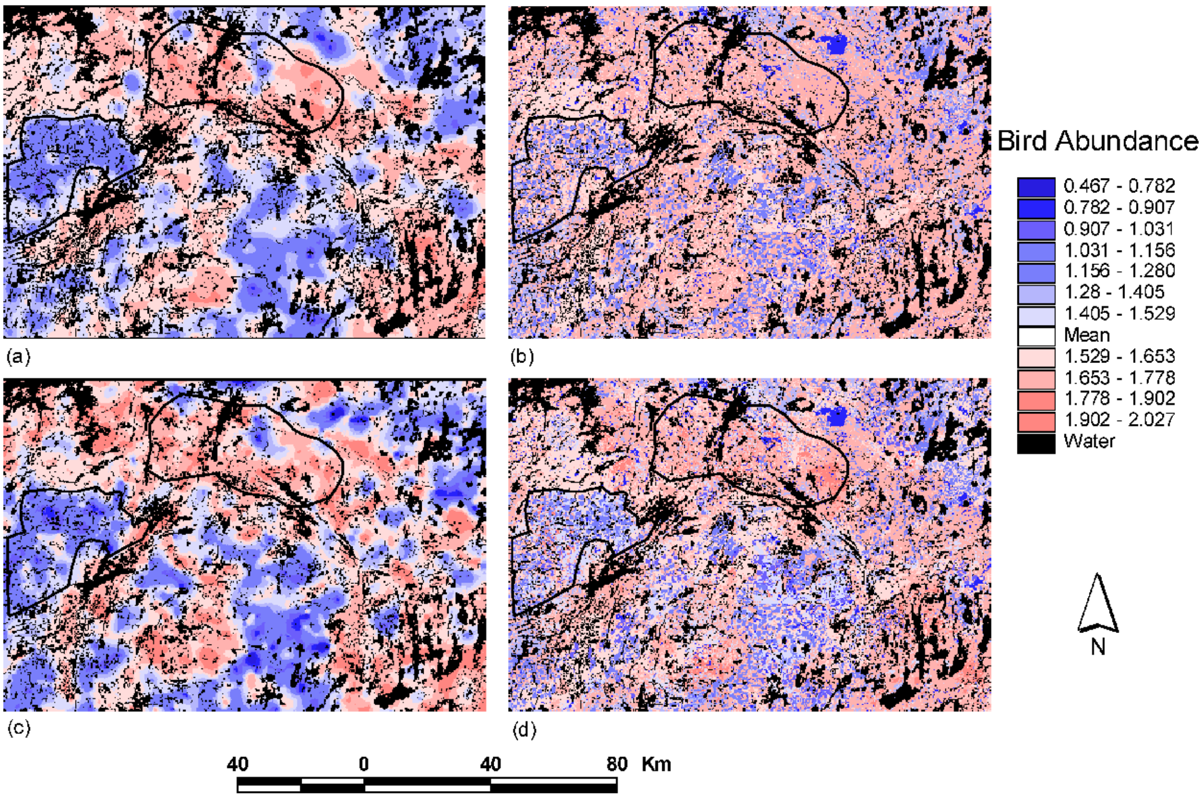

Figure 6. Comparison of maps created from (a) global OPK, (b) non-spatial habitat-based assignment, (c) co-kriging, and (d) stratified OPK. Polygon on mid-left side delineates the fragmented landscape, and polygon at upper-middle delineates the un-fragmented landscape. Colors as in Figure 1c.

Table 5. Comparison of mean bird density values (in mature forest only) between reference map and maps created from alternative mapping approaches.

Landscape[1]	Reference value	Habitat alone	Spatial alone (global OPK)	Spatial (co-kriging)	Spatial + habitat (stratified OPK)
Fragmented	1.5168	1.6191	1.1989	1.2717	1.5686
Un-fragmented	1.6953	1.6717	1.5945	1.6512	1.6998

[1]Fragmented refers to age-class fragmentation created by dispersed harvest-block cutting.

be achieved by utilizing a GIS to display a coarse satellite-based landcover map and road network, and to set sample clusters 5 km apart.

The landcover classification was simple, and included categories such as mature versus young conifer or mixedwood. Points within clusters were constrained to be a minimum distance apart (e.g., 250 m), and to be at least 100 m from stand boundary. Approximate locations of sample points were uploaded from the GIS into a hand-held GPS, and field staff using the GPS recorded the exact location of the sample sites, after adjusting the site location based on stand-edge proximity. The nearest-waypoint function of the hand-held GPS was used to insure the proper distance between sample points was maintained. In this study, three field technicians sampled 2 points within each cluster (totaling 6 points per cluster), and generally 3 clusters (18 points) were sampled before 10 am, when the songbird recording-window closes.

Although the assumption of randomness is partially violated (as it almost always is for boreal forest sampling because of access issues), the spatial dispersion

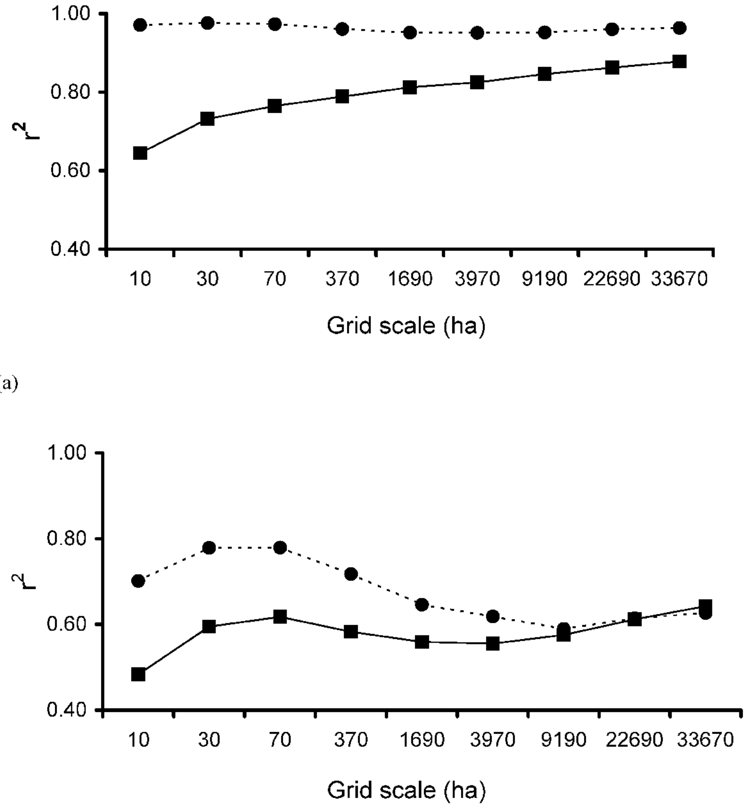

Figure 7. Multi-scale performance assessment of stratified (circles) versus global (squares) ordinary point kriging (OPK) for low variance (a) and high variance (b) maps, respectively; r^2 is the correspondence of bird density (weighted means) in hexagons for the reference and interpolated maps, respectively. For the low variance map (a), performance of stratified OPK is consistently higher than global OPK, but performance for the two methods converge at broader spatial scales.

of sample points is suitable for estimating spatial-autocorrelation, and habitats are sampled in proportion to their occurrence at the local (2500 ha) scale.

Discussion

The results of this study provide some insight into the performance of alternative sampling and mapping approaches for reconstructing patterns of distribution and abundance. The most obvious concerns and limitations of the study are that (1) the results pertain to a simulated bird species, not a real population, (2) the range of simulated site-level variation for LV1, LV2, HV1 and HV2 may not adequately reflect the range of site-level variation for real boreal birds, (3) the simulated species adhere to relatively well defined habitat associations, and real species may not, (4) habitat association mean values are stable across the study area, and (5) there is no variation in detection probability at any point. However, this study is focused at understanding the relative performance of alternative sampling and interpolation techniques, so more realistic simulation of regional level variation in ecological behavior of birds would likely not have changed any of the mapping performance interpretations. An alternative approach would be to simulate multiple reference landscapes using a general model, such as that developed by Saura and Martinez-Millan (2000). This would permit a more general analysis of sampling and interpolation performance, but would be less realistic than our approach. A useful next step would be to examine the effectiveness of the mixed-cluster sampling approach under a variety of landscape patterns and spatial distributions that are simulated for different ecoregions, but still based on satellite-interpreted landcover maps.

The most successful sampling technique in this study was the mixed-cluster strategy. This approach was a combination of systematic, random, and clus-

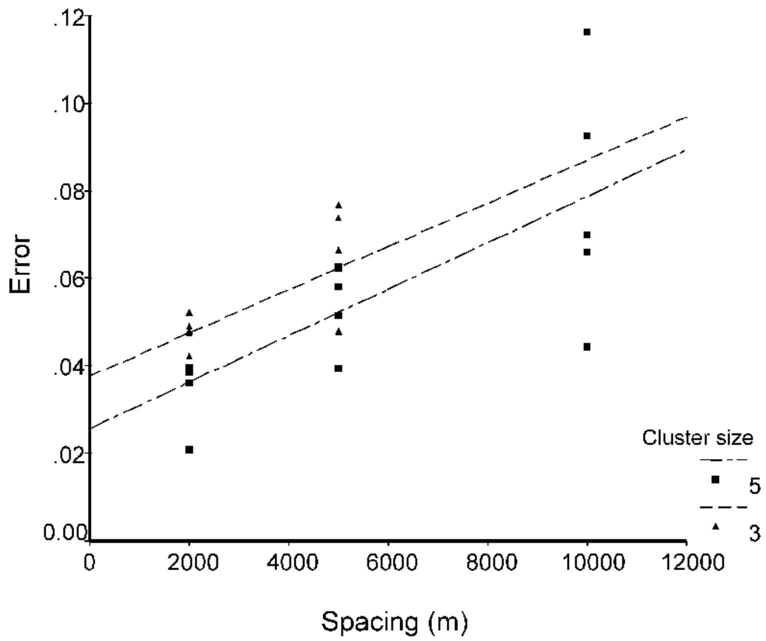

Figure 8. Effect of cluster spacing on error (distance) between weighted-mean values for reference and interpolated maps in the fragmented landscape (see Figure 6d). Interpolated map was created through stratified OPK using mixed-cluster point sampling data. The accuracy of bird density estimates derived from the sampling scheme of 5-point clusters at 10,000 m spacing was poor relative to the 5,000 and 2,000 m spacing scheme.

tered components, in that the selection starting point is regularly spaced (the centroid of the hexagonal grid), and all selected points were random locations, within a specified distance from the starting point. The procedure was stratified in two ways: water was excluded from the samples, and also there was a contextual constraint of restricting samples to patches of $>$ 10 ha, and $>$ 100 m from edge. This ensured that the sample point is placed in a predominant habitat type found in the vicinity of the starting point. This is a complex sampling strategy that appears to effectively capture the underlying spatial variance that is characteristic of a forested landscape shaped by contagious and spatially structured events such as wildfire and logging. Other sampling strategies may perform better in other types of landscapes, so prestudy modeling and evaluation of strategies is an important first step. Fortin et al. (1989) also found that sample designs that incorporated varying sample steps, such as the "systematic-cluster design", were most suitable for detecting spatial structure.

Ordinary point kriging has become a popular approach to mapping patterns of distribution and abundance in animals (e.g., Villard and Maurer 1996; McKenney et al. 1998; Venier et al. 1999) but it is not the only approach to interpolation, nor always the best. When kriging is applied to all sample points, the model is termed global kriging (Wallerman et al. 2002). For terrestrial ecology, a weakness of global kriging is the inability to recognize discrete strata that separate the functional response of animals. Consider, for example, a bird that prefers to nest in the canopy of tall trees. Interpolation of point samples that fall on the opposite sides of a clear cut will not account for this discontinuity in habitat. In geostatistical terms, the spatial discontinuity results in a statistical model that violates the principle of stationarity (Goovaerts 1997), and kriging within strata (stratified OPK) becomes a reasonable alternative (e.g., Fortin et al. 1989; Goovaerts 1997; Gunnarsson et al. 1998; Burrough 2001). Of the 4 approaches that we investigated to address bird abundance mapping, only stratified OPK was able to effectively detect both the spatial and habitat relationships found within the data.

Assessing the performance of global OPK, including OPK with co-variates, can be done relatively easily using cross-validation statistics. Assessing performance of stratified OPK is more complex. On one hand, performance measures are improved because habitat information is incorporated by the stratifica-

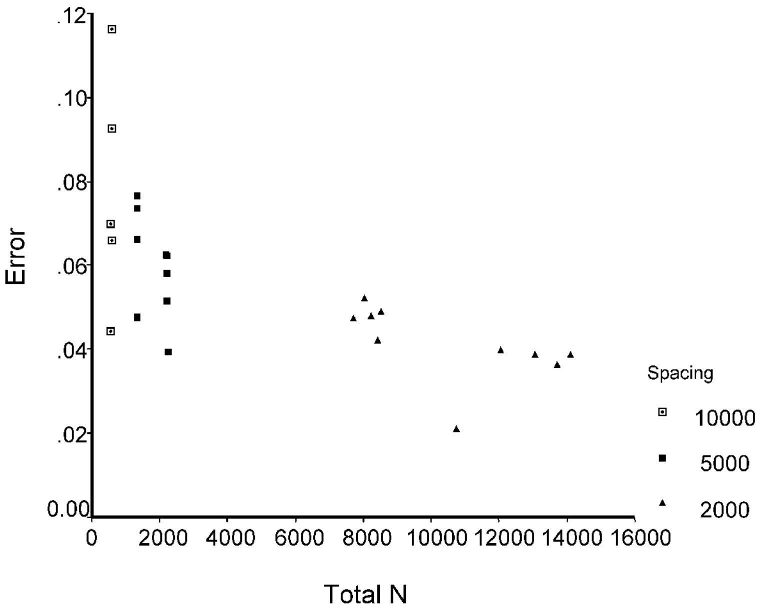

Figure 9. Effect of total sample size on error (distance) between weighted mean values for reference and interpolated maps in the fragmented landscape. The error of bird density estimates derived from the sampling scheme of clusters at 5,000 m spacing (square symbols) was slightly higher, and more variable, than for schemes with 2,000 m spacing (triangle symbols) schemes. Error for 10,000 m spacing was unacceptably high.

tion procedure, but on the other hand, the lower number of point samples within a stratum decreases performance statistics. One solution that deserves further investigation is to combine the separate experimental semivariograms into a pooled semivariogram, where the pooled value is weighted by the number of pairs in the relevant contributing semivariogram (Goovaerts 1997).

The accumulation and interaction of errors across and among layers is more complex than a simple joint probability of errors derived from the individual ordinary point kriging layers. The spatial variance is not stationary, and the validity of creating a stratified-kriging variance map by overlaying the contributing OPK variance maps is questionable. As well, the reliability of the estimates is in part a function of sampling strategy. Our approach to assessing error in this study was empirical (we used Monte Carlo techniques to compare simulated densities in two landscapes in a manner that would resemble an actual application of the data), but further theoretical work may be helpful (e.g., Sadahiro 1999). None-the-less, the use of simulated abundance data based on actual landcover data, with bird means and variances reflecting dominant habitat patterns on the landscape, is an important first step in understanding the influence of various sampling and mapping strategies on the ability to detect landscape-level effects on animal abundance. With stratified OPK, the habitat value of every cell is known; hence finite sampling theory may be applicable for estimation and inference, and this also deserves further investigation (see Valliant et al. 2000).

This study of computer-generated sampling schemes enabled the evaluation of general principles for alternative sampling schemes, as well as detailed parameter selection. We found that implementing the mixed-cluster sampling scheme in the boreal forest involved incorporating general principles into actual field sample-point selection, but we could not use exact locations provided through computer generation of sample points. In general terms, the mixed-cluster scheme ensures a relatively even but staggered distribution of clusters across the landscape. Incorporating a random component to the starting point of the cluster ensures variability in inter-cluster distances, and incorporating random variability to the within-cluster point locations ensure habitats are selected approximately in proportion to their occurrence at a local scale, and ensures added variability in inter-point distances. This spectrum of inter-point distances improves the estimation of spatial auto-correlation, and the relatively even dispersion of clusters across the landscape reduces the effects of spatial holes in the abundance estimates. Mapped spatial error (variance) decreases as a function of the number of sample points available in the neighborhood window, thus a relatively even dispersion of points is necessary to maintain relatively homogeneous variance across the interpolated map. In logistical terms, the clustered approach is one of the most cost-effect techniques for large-scale avian monitoring (Carlson 2001), and if modified appropriately, is also very effective for estimating spatial auto-correlation and producing maps (spatially explicit models) of wildlife abundance. Future enhancements to the computer-sampling program will include distance-related cost-functions to permit more accurate cost-effectiveness evaluation of alternative sampling schemes.

Conclusions

Within the scope and limitations of this study (as discussed above) the following conclusions are made about sampling performance for the objective of medium to broad scale mapping of avian abundance.

1. Stratified OPK was the most effective of the analysis techniques investigated. It successfully detects and models both spatial and non-spatial (stand level habitat association) information from the dataset to reconstruct patterns of distribution and abundance. This was the only technique to effectively discriminate bird density in landscapes where a spatial fragmentation effect had been encoded in the reference data set.
2. Performance of ordinary and stratified OPK tends to converge when evaluated at broad spatial scales $>$ 4000 ha. OPK can be used effectively for broad scale, regional analysis, but stratified OPK is more effective for finer scale, habitat based analyses.
3. Mixed-cluster sampling, as uniquely defined in this study, is an effective sampling strategy. Bias in the estimate of abundance was low, mapping performance across spatial scales was high, and mixed-cluster sampling performed equally well to alternative methods in terms of modeling structural variance. Logistically, it is among the easiest and least costly strategies to implement in the field.
4. Mapping performance increases significantly if sampling is restricted to habitat patches $>$ 10 ha. Restricting sample points to locations $>$ 100 m from stand edge produced similar results because stands must be at least 9 ha to have potential sample point locations beyond that minimum distance.
5. For mixed-cluster sampling, the errors in both global and stratified OPK are first a function of the spacing between clusters, and second, the number of points within a cluster. For this simulation study, both 3-point clusters at 2000 m staggered spacing and 5-point clusters at 5000 m staggered spacing had sufficient "spatial power" to correctly detect landscape-level effects of fragmentation if the maps were created using stratified OPK. This corresponds to about 8000 and 2000 points, respectively, or 0.8 and 0.2 points / km^2.
6. In practice, a close approximation of the computer generated mixed-cluster sampling was achieved by utilizing the combined spatial data technologies of satellite imagery, road network maps, GIS and GPS spatial data management software, and hand-held GPS units. Sample clusters were set 5000 m apart, and points within clusters were constrained to be a minimum of 250 m apart, and to be at least 100 m from stand boundary or road.

Acknowledgements

Very kind thanks are extended to Julie Elliott, who ran most of the computerized sampling and kriging analyses, and who organized and summarized the mass of data output. Many thanks also for discussion with Marie-Josee Fortin, and especially her thoughts on stratified kriging and cluster sampling. Reviewers of the manuscript provided very constructive comments that improved the final paper.

References

Andren H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos 71: 355–366.

Austen M.J.W., Francis C.M., Burke D.M. and Bradstreet M.S.W. 2001. Landscape context and fragmentation effects on forest birds in southern Ontario. Condor 103: 701–714.

Bellehumeur C. and Legendre P. 1998. Multiscale sources of variation in ecological variables: modeling spatial dispersion, elaborating sampling designs. Landscape Ecology 13: 15–25.

Benedetti-Cecchi L. 2001. Beyond Baci: Optimization of environmental sampling designs through monitoring and simulation. Ecological Applications 11: 783–799.

Boone R.B. and Krohn W.B. 2000. Predicting broad-scale occurrences of vertebrates in patchy landscapes. Landscape Ecology 15: 63–74.

Boulinier T., Nichols J.D., Hines J.E., Sauer J.R., Flather C.H. and Pollock K.H. 2001. Forest fragmentation and bird community dynamics: Inference at regional scales. Ecology 82: 1159–1169.

Burrough P.A. 2001. GIS and geostatistics: Essential partners for spatial analysis. Environmental and Ecological Statistics 8: 361–377.

Carlson M. 2001. Cost-effective sampling design for large-scale avian monitoring. MS Thesis, University of Alberta: Edmonton, Alberta, Canada. 89 pp.

Dale M.R.T. and Fortin M.J. 2002. Spatial autocorrelation and statistical tests in ecology. Ecoscience 9: 162–167.

Davis B.M. 1987. Uses and abuses of cross validation in geostatistics. Mathematical Geology 19: 242–248.

Dessard H. 1999. Comparison of design-based and model-based estimates for tropical forestry resource with post-stratification. Annals of Forest Science 56: 651–665.

Elkie P.C. and Rempel R.S. 2001. Detecting scales of pattern in boreal forest landscapes. Forest Ecology and Management 147: 253–261.

Elliott J.M., 1977. Some Methods for the Statistical Analysis of Samples of Benthic Invertebrates. The Ferry House, England.

Fortin M.J., Drapeau P. and Legendre P. 1989. Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83: 209–222.

Fortin M.J. and Payette S. 2002. How to test the significance of the relation between spatially autocorrelated data at the landscape scale: A case study using fire and forest maps. Ecoscience 9: 213–218.

Foster D.R., Knight D.H. and Franklin J.F. 1998. Landscape patterns and legacies resulting from large, infrequent forest disturbances. Ecosystems 1: 497–510.

Foster J.R. 2001. Statistical power in forest monitoring. Forest Ecology and Management 151: 211–222.

Goovaerts P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, New York, USA.

Gunnarsson F., Holm S., Holmgren P. and Thuresson T. 1998. On the potential of Kriging for forest management planning. Scandinavian Journal of Forest Research 13: 237–245.

Hall F.G., Strebel D.E. and Sellers P.J. 1988. Linking knowledge among spatial and temporal scales: vegetation, atmosphere, climate and remote sensing. Landscape Ecology 2: 3–22.

Hayes J.P. and Steidl R.J. 1997. Statistical power analysis and amphibian population trends. Conservation Biology 11: 273–275.

Hobson K.A. and Schieck J. 1999. Changes in bird communities in boreal mixedwood forest: Harvest and wildfire effects over 30 years. Ecological Applications 9: 849–863.

Hobson K.A., Kirk D.A. and Smith A.R. 2000. A multivariate analysis of breeding bird species of western and central Canadian boreal forests: Stand and spatial effects. Ecoscience 7: 267–280.

Hunsaker C.T., Goodchild M.F., Friedl M.A. and Case T.J. 2001. Spatial Uncertainty in Ecology. Springer-Verlag New York, Inc.: New York, New York, USA.

Kotliar N.B. and Wiens J.A. 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59: 253–260.

Legendre P. and Fortin M.J. 1989. Spatial pattern and ecological analysis. Vegetatio 80: 107–138.

Legendre P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74: 1659–1673.

Lichstein J.W., Simons T.R., Shriner S.A. and Franzreb K.E. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecological Monographs 72: 445–463.

Liebhold A.M. and Gurevitch J. 2002. Integrating the statistical analysis of spatial data in ecology. Ecography 25: 553–557 .

Lin Y.P. and Rouhani S. 2001. Multiple-point variance analysis for optimal adjustment of a monitoring network. Environmental Monitoring and Assessment 69: 239–266.

Lowell K. and Jaton A. 1999. Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources. Ann Arbor Press, Chelsea, Michigan, USA.

McKenney D.W., Rempel R.S., Venier L.A., Wang Y.H. and Bisset A.R. 1998. Development and application of a spatially explicit moose population model. Canadian Journal of Zoology 76: 1922–1931.

Meisel J.E. and Turner M.G. 1998. Scale detection in real and artificial landscapes using semivariance analysis. Landscape Ecology 13: 347–362.

Potvin F., Lowell K., Fortin M.J. and Belanger L. 2001. How to test habitat selection at the home range scale: A resampling random windows technique. Ecoscience 8: 399–406.

Rossi R.E., Mulla D.J., Journel A.G. and Franz E.H. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62: 277–314.

Roy P.S. and Tomar S. 2000. Biodiversity characterization at landscape level using geospatial modelling technique. Biological Conservation 95: 95–109.

Royle J.A. and Berliner L.M. 1999. A hierarchical approach to multivariate spatial modeling and prediction. Journal of Agricultural Biological and Environmental Statistics 4: 29–56.

Ryan K.C. 2002. Dynamic interactions between forest structure and fire behavior in boreal ecosystems. Silva Fennica 36: 13–39.

Sadahiro Y. 1999. Accuracy of areal interpolation: a comparison of alternative methods. CSIS Discussion Paper No. 9. Center for Spatial Information Science and Department of Urban Engineering, University of Tokyo, Tokyo, Japan, 41 p.

Saura S. and Martinez-Millan J. 2000. Landscape patterns simulation with a modified random clusters method. Landscape Ecology 15: 661–678.

Steidl R.J., Hayes J.P. and Schauber E. 1997. Statistical power analysis in wildlife research. Journal of Wildlife Management 61: 270–279.

Thompson F.R., Burhans D.E. and Root B. 2002. Effects of point count protocol on bird abundance and variability estimates and power to detect population trends. Journal of Field Ornithology 73: 141–150.

Valliant R., Dorfman A.H. and Royall R.M. 2000. Finite Population Sampling and Inference: A Prediction Approach. John Wiley & Sons Inc., New York, New York, USA..

van Groenigen J.W. 2000 . The influence of variogram parameters on optimal sampling schemes for mapping by kriging. Geoderma 97: 223–236.

Venier L.A., McKenney D.W., Wang Y. and McKee J. 1999. Models of large-scale breeding-bird distribution as a function of macro-climate in Ontario, Canada. Journal of Biogeography 26: 315–328.

Vernier P.R., Schmiegelow F.K.A., and Cumming S.G. 2002. Modelling bird abundance from forest inventory data in the boreal mixedwood forests of Canada. In: Scott J.M., Heglund P.J. and

Morrison M.L., Predicting Species Occurrences: Issues of Accuracy and Scale, pp. 559–572. Island Press: Washinton, D.C., USA.

Villard M.-A. and Maurer B.A. 1996. Geostatistics as a tool for examining hypothesized declines in migratory songbirds. Ecology 77: 59–68.

Wallerman J., Joyce S., Vencatasawmy C.P. and Olsson H. 2002. Prediction of forest stem volume using kriging adapted to detected edges. Canadian Journal of Forest Research 32: 509–518.

Zar J.H. 1984. Biostatistical Analysis. Prentice Hall, Englewood Cliffs, New Jersey, USA.