

RESEARCH ARTICLE

Accommodating temporary emigration in spatial distance sampling models

Jeremy D. Mizel¹  | Joshua H. Schmidt²  | Mark S. Lindberg³¹Arctic Network, U.S. National Park Service, Fairbanks, AK, USA²Central Alaska Network, U.S. National Park Service, Fairbanks, AK, USA³Department of Biology and Wildlife Ecology, University of Alaska Fairbanks, Fairbanks, AK, USA**Correspondence**Jeremy D. Mizel
Email: jeremy_mizel@nps.gov**Funding information**

U.S. National Park Service's Inventory and Monitoring Program and the Arctic Network

Handling Editor: Jacqueline Frair

Abstract

1. Model-based distance sampling is commonly used to understand spatial variation in the density of wildlife species. The standard approach assumes that individuals are distributed uniformly and models spatial variation in density using plot-level effects. Thinned point process (TPP) models for surveys of unmarked populations (spatial distance sampling) better leverage the spatial information underlying individual encounters, and in the presence of within-plot variation in density, may explain a larger proportion of the spatial variation in density. However, existing spatial distance sampling approaches are conditioned on the assumption that all individuals are present and available for sampling. Temporary emigration of individuals can therefore result in biased estimates of abundance.
2. We extended spatial distance sampling models to accommodate temporary emigration (TPP model). Using simulations of a thinned inhomogeneous point process, we assessed the performance of the TPP model relative to the temporary emigration distance sampling (TEDS) model, which implies a uniform distribution of individuals. In addition, we compared inferences between TPP and TEDS models using data for two passerine species in Alaska.
3. Parameter estimates from the TPP model exhibited improved coverage probability and precision relative to the TEDS model including a 26% reduction in the coefficient of variation (CV) of the population size estimate.
4. In the applied example, the TEDS model indicated weak relationships between abundance and habitat covariates, whereas the TPP model indicated strong relationships for those same effects, suggesting that spatial distance sampling models can provide considerably stronger inference in the presence of within-plot variation in density. In addition, the CV of the population size estimates for the two passerine species were 32% and 4% smaller under the TPP model.
5. *Synthesis and applications.* We expect our extension accommodating temporary emigration will be a critical specification for spatial distance sampling models, particularly for studies assessing changes in the distribution and abundance of highly mobile species including passerines.

KEYWORDS

Bayesian hierarchical models, inhomogeneous point process, long-term monitoring, N-mixture model, over-dispersion, spatial distance sampling, spatially explicit models, superpopulation, temporary emigration

1 | INTRODUCTION

Conventional and model-based distance sampling are commonly used to estimate the density of wildlife species (Buckland et al., 2001, 2004). However, standard distance sampling methods do not fully leverage the spatial information underlying individual encounters for use in explaining variation in density. Standard model-based approaches rely on plot-level effects for explaining spatial variation in density, inducing over-dispersion when density varies within plots. In contrast, spatial distance sampling describes the observed locations of individuals as arising from a spatial point process thinned through incomplete detection and, consequently, exploits the spatial information inherent in the location of encounter for potentially stronger inference about spatial processes (Hedley & Buckland, 2004; Johnson, Laake, & Ver Hoef, 2010; Kéry & Royle, 2016; Yuan et al., 2016).

However, existing spatial distance sampling approaches describe the observation process entirely in terms of the detection function (i.e. the decline in detectability as a function of individual distance from the transect) and, consequently, are conditioned on the assumption that all individuals are present and available for sampling. The limitations of this approach are best understood by decomposing the observation process into its four primary components: (1) the probability the individual's home range overlaps the sampling unit, p_s ; (2) the probability the individual is present within the area that is exposed to sampling during the survey occasion, p_p ; (3) the probability the individual is available for detection during the survey occasion, p_a ; and (4) the probability that the individual is detected given that it is present and available, p_d (see Nichols, Thomas, & Conn, 2009).

The various field methods used in wildlife surveys differ in the components of the detection process that they estimate and, consequently, provide inference to different subsets of the superpopulation; i.e. the population comprised of all individuals with a non-negligible probability of being present in the sample unit over the survey period (Kendall, Nichols, & Hines, 1997). Single-visit distance sampling typically only addresses p_s (through the survey design) and p_d (through the detection function). By failing to address p_p and p_a directly, single-visit approaches make the implicit assumption that these components of the detection process are relatively constant through space and time. Should this assumption prove untenable, then inference is limited to some unknown proportion of the superpopulation and abundance estimates are negatively biased (Chandler, Royle, & King, 2011; Nichols et al., 2009; Schmidt, McIntyre, & MacCluskie, 2013). Thus, estimates of superpopulation abundance require fewer implicit assumptions and, therefore, should have greater sensitivity for assessing population and spatial dynamics.

Chandler et al. (2011) developed a model framework for explicitly accommodating non-Markovian temporary emigration in unmarked populations. Their model is applicable to a number of protocols including the combination of repeat surveys and distance sampling which allows separate estimation of p_d and the composite probability of availability $p_p p_a$. However, the temporary emigration distance sampling (TEDS) version of their model implies that individuals are distributed uniformly in space and models spatial variation in abundance at

the plot-level. Models with plot-level effects (e.g. the TEDS model) will often be a good description of variation in abundance when sample units are small or habitat gradients are gradual (Kéry & Royle, 2016; Miller, Burt, Rexstad, & Thomas, 2013). However, an inhomogeneous distribution of individuals over a compressed habitat gradient would be expected to induce considerable within-plot variation in density. Under these conditions, properly specifying the distribution of individuals within plots as inhomogeneous may explain a larger proportion of the variation in density.

There have been several recent developments in spatial distance sampling inference. Using full likelihood approaches (i.e. the N_j as explicit parameters), Johnson et al. (2010) and Yuan et al. (2016) developed thinned point process models for continuous space. Kéry and Royle (2016) formulated a thinned point process model for discrete space using a data augmentation approach. While these methods allow fine-scale assessment of spatial variation in abundance, they do not address the temporary emigration process and, consequently, may yield biased estimates of abundance.

Herein, we unify the thinned point process formulation of the Kéry and Royle (2016) model and the full likelihood description of the temporary emigration process developed for N-mixture models (Chandler et al., 2011). The development of these inference methods was motivated by a need to efficiently and effectively monitor arctic passerine populations, which are expected to undergo rapid changes in abundance and distribution in association with shrub expansion into open habitats and higher elevations (Dial et al., 2015; Mizel, Schmidt, McIntyre, & Roland, 2016; Sturm, Racine, & Tape, 2001; Tape, Sturm, & Racine, 2006). Thus, it was important that we develop inference methods with the sensitivity to efficiently detect and distinguish population and spatial dynamics while also addressing all components of the detection process. We viewed accommodating temporary emigration as critical given the characteristics of sampling arctic passerine populations including territorial movements of individuals across plot boundaries between survey visits (i.e. spatial temporary emigration) and large variation in cue production over a highly compressed breeding season (i.e. random temporary emigration) (Kéry & Royle, 2016). While motivated by our particular case study, the resulting framework will be broadly applicable in other situations where temporary emigration occurs. Our specific objectives in this manuscript were to: (1) develop a framework for accommodating temporary emigration in a spatial distance sampling context; (2) assess the precision and bias of estimates from our model relative to the TEDS model through simulation and (3) demonstrate the advantages of our unified model through an application to field data for two passerine species.

2 | MATERIALS AND METHODS

2.1 | Model development

Herein, we develop the temporary emigration, thinned point process (hereafter, TPP) model for a hierarchical sampling design where multiple line or point transect plots are each comprised of a grid of pixel centroids from an underlying habitat raster, and each plot is surveyed

multiple times (e.g. separated by a day or a few days). The survey yields the locations of the observed individuals in continuous space, but these are subsequently assigned to the nearest pixel centroid for the data structure used in the analysis. For simplicity of presentation, we develop the model in the context of a single season.

Following Kéry and Royle (2016), individuals are distributed according to a Poisson point process with intensity function λ . However, in order to accommodate a replicated count protocol and temporary emigration in the generalized N-mixture framework (i.e. Chandler et al., 2011), we use a full likelihood formulation with a conditional multinomial observational model (Kéry & Royle, 2016; Royle, Dawson, & Bates, 2004). Data from the replicated count-spatial distance sampling protocol are organized as a series of counts of unmarked individuals within each pixel $i = 1, 2, \dots, I$, in plot $j = 1, 2, \dots, J$, and visit $k = 1, 2, \dots, K$. This is amenable to formulating the observation model in terms of multinomial probabilities and leads to an implicit binning of distances (Kéry & Royle, 2016). The focus is on estimating the superpopulation point intensity function for a Poisson point process of the individual locations (Kéry & Royle, 2016):

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 x_{ij}, \quad (1)$$

where λ_{ij} is the superpopulation intensity function for each pixel, x_{ij} is the pixel covariate value, and β_0 and β_1 are parameters to be estimated. Integrating over the grid of points in each plot yields the expected plot-level superpopulation sizes Λ_j (Kéry & Royle, 2016). By the rectangular rule, this integral is approximated for the discretized region using summation (Kéry & Royle, 2016):

$$\Lambda_j = \sum_i \exp(\beta_0 + \beta_1 x_{ij}) \quad (2)$$

The local superpopulation sizes are specified as Poisson random variables:

$$M_j \sim \text{Poisson}(\Lambda_j).$$

The populations available for detection on a given plot-visit are binomial outcomes of the superpopulation sizes and the probability of availability ϕ (i.e. the product of p_p and p_d) which is the complement of the temporary emigration probability (Chandler et al., 2011): $N_{j,k} \sim \text{Binomial}(M_j, \phi)$. Models for time- and plot-dependent ϕ could also be considered including random effects parameterizations which account for over-dispersion and/or temporal correlation in repeated counts (Buckland, Oedekoven, & Borchers, 2016). Here, we assume ϕ is constant for simplicity of presentation.

The hierarchical distance sampling (HDS) model of Royle et al. (2004) and other standard model-based formulations (Buckland et al., 2016) imply that the within-plot distribution of individuals is uniform. In the HDS model, the observed counts follow a multinomial distribution where the vector of cell probabilities is the product of detection probability and of occurring in a particular distance class, which is specified as uniform across distance classes (adjusted for annulus area in the point transect case). In the TPP model, the multinomial observation probabilities are the product of detection probability and the probability of occurring in a particular pixel θ_{ij} . The latter is derived

using the point intensity function for each pixel, conditional on the plot-specific total (Kéry & Royle, 2016):

$$\theta_{ij} = \frac{\lambda_{ij}}{\sum_i \lambda_{ij}}. \quad (3)$$

Using a half-normal detection function as an example, the multinomial observation cell probabilities are:

$$\pi_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right) * \theta_{ij}, \quad (4)$$

where d is the distance from the transect to the pixel centroid and σ is the scale parameter of the detection function. The probability of not being detected at all is $\pi_{0j} = 1 - \sum_i \pi_{ij}$ and the conditional multinomial cell probabilities are (Kéry & Royle, 2016; Royle et al., 2004):

$$\pi_{ij}^c = \frac{\pi_{ij}}{1 - \pi_{0j}}. \quad (5)$$

The model for the observations is decomposed into multinomial and binomial components in order to avoid specifying the multinomial index as a random variable (Kéry & Royle, 2016; Royle et al., 2004). Conditional on the observed count, the number of individuals observed in a pixel on each plot-visit follows a multinomial distribution. The observed counts $n_{j,k}$ are binomial outcomes of the latent $N_{j,k}$ and the total probability of detection $1 - \pi_{0j}$:

$$y_{i,j,k} | n_{j,k} \sim \text{Multinomial}(n_{j,k}, \pi_{ij}^c)$$

$$n_{j,k} \sim \text{Binomial}(N_{j,k}, 1 - \pi_{0j}).$$

2.2 | Simulation study

We designed our simulations in the context of a passerine bird monitoring in an open arctic habitat. We began by simulating a spatially correlated covariate (i.e. a continuous measure of habitat) on a grid of 10×10 pixels, replicated over 150 plots under a line transect sampling design (Figure 1). Each plot-specific vector of covariate values x was generated as a series of multivariate random variables using the Cholesky decomposition of the variance-covariance matrix $V^{1/2}$ and its product with a compatible vector of iid normal random variables $e \sim N(0,1)$ (Kéry & Royle, 2016). We simulated point intensity for each discretized plot as $\log(\lambda_{ij}) = \beta_0 + \beta_1 x_{ij}$ where $\beta_0 = -5.8$ and $\beta_1 = 2$, corresponding to an inhomogeneous point process. The local superpopulation sizes were generated from a Poisson distribution with an expected value equal to $\sum_i \exp(\beta_0 + \beta_1 x_{ij})$ (Kéry & Royle, 2016). We used the plot-specific, conditional intensity probabilities to allocate individuals to pixels for each plot-visit. A vertical line transect bisected the centre of each plot resulting in a maximum distance of 4.5 to any pixel centroid (Figure 1). Observations were simulated over four replicated counts using $\phi = 0.6$ and a half-normal detection function with $\sigma = 3$.

We fit the TPP model to the simulated data to assess whether it could recover the data-generation parameter values. We also investigated model performance relative to the TEDS model (Chandler et al., 2011; Kéry & Royle, 2016). Observations for the TEDS model were

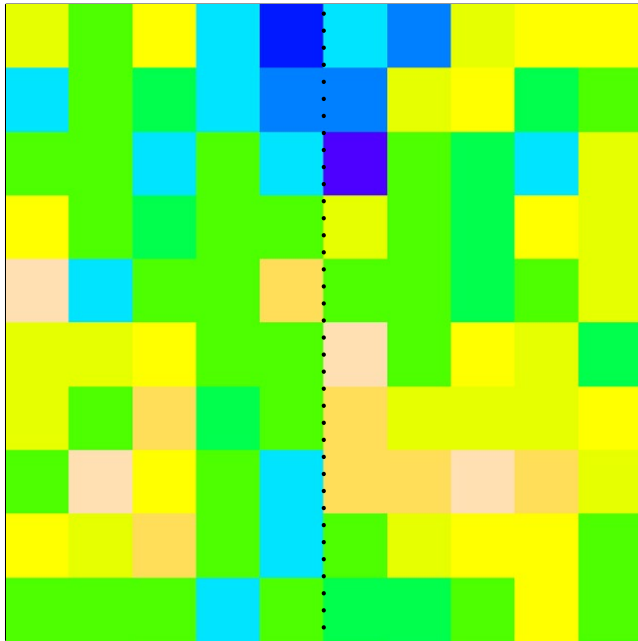


FIGURE 1 A spatially correlated covariate simulated on a 10×10 pixel grid. Each plot in the simulation study was bisected by a line transect (dashed line)

generated as above (i.e. the TPP model was the data generating model) except that data structure was reduced to the number of individuals for each plot-visit in distance class $i = 1, 2, \dots, 5$ rather than pixel $i = 1, 2, \dots, 100$. We used five distance classes corresponding to the sequence of centroid distances in the discretized plots, paralleling the implicit binning of the TPP model. We used the conditional multinomial formulation of the TEDS model (Kéry & Royle, 2016) making it structurally similar to the TPP model. However, the TEDS model assumes a uniform distribution of individuals across distance classes. The TEDS model is made spatially explicit by a model for variation in local super-population abundance, which we specified as $\log(\lambda_j) = \beta_0 + \beta_1 \bar{x}_j$, where $\bar{x}_j$ is the mean covariate value over all grid points in a plot. Because this represented a reduced information summary of the data generating process, we assumed the presence of unexplained variation and modelled this by including a random effect in the log-transform of λ_j . Thus, we included both Poisson and Poisson log-normal versions of the TEDS model in comparison to the TPP model.

We ran 500 simulations for each of the three models. To improve mixing of chains in the TPP and TEDS models, we removed the binomial model for the N_{jk} and described the observations a binomial outcomes with M_j sample size and probability $(1 - \pi_{0j}) * \phi$ (Kéry & Royle, 2016). We fit all models in a Bayesian framework using JAGS version 4.0.0 (Plummer, 2003) via the jagsUI package (Kellner, 2016) in program R 3.2.2 (R Core Team, 2015). We specified vague normal priors for all regression coefficients, a uniform prior for σ from 0 to the plot truncation distance, and a uniform (0,1) prior for ϕ . We used uniform (0,100) priors for the standard deviations of the random effects. Summaries of the posterior distribution were calculated from one Markov chain run for 21,500 iterations with a 6,500 iteration burn-in and thinning every five draws with the exception of the log-normal

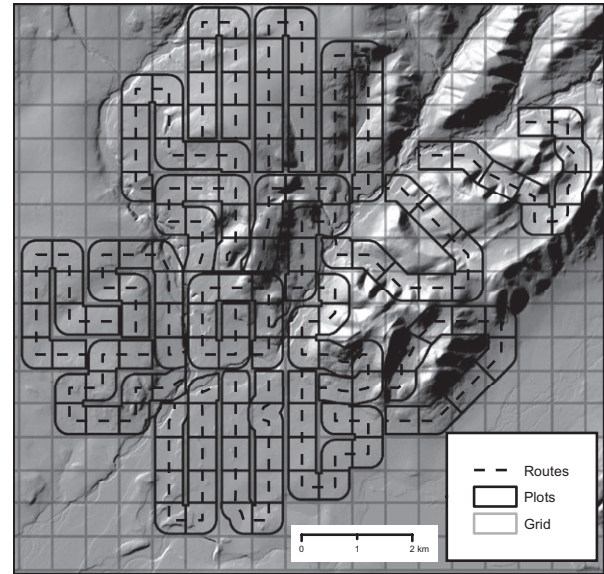


FIGURE 2 The sampling design implemented at a study site in Noatak National Preserve, Alaska in 2016. Plots (solid borders) and routes (dashed lines) are shown for a species with a 250 m truncation distance

TEDS model, which we ran for 58,000 iterations with a 28,000 iteration burn-in. We report the average parameter estimates, average relative bias, root mean squared error (RSME) and coverage (percentage of 95% Bayesian credible intervals for parameters that overlap the true values). See Appendices S1–S2 for simulation and JAGS model code.

2.3 | Application: Arctic passerines

We applied the TPP model to data collected in the Noatak National Preserve, Alaska between May 21 and June 19 2016 for two species of passerines, American tree sparrow (*Spizella arborea*) and savannah sparrow (*Passerculus sandwichensis*). Our study site represents one of three long-term, landbird monitoring sites in the U.S. National Park Service's Arctic Inventory and Monitoring Network and covers an elevational gradient from lowland to alpine habitats with open and riparian shrublands, *Dryas* and/or mixed dwarf shrub tundra, and barren slopes. Habitat types are distributed non-uniformly, presumably influencing the distribution of bird territories. We adopted a route-based, continuous-effort survey design to maximize encounter rates (i.e. exposure to sampling) and study area coverage. We centred a 10.8×9.6 km grid comprised of $600 \text{ m} \times 600 \text{ m}$ cells on the airstrip at which we located our basecamp. We established fixed survey routes that were designed to be completed in a single day and changed direction (Hiby & Krishna, 2001) in part to return observers close to their original starting point (Figure 2). In general, routes were located such that they intersected the centroid of each cell and then were deflected at an angle of 0 or 90 degrees (i.e. followed a parallel or perpendicular direction; Figure 2). In some cases, we shifted routes off of the central axis of the cell due to the presence of terrain features (e.g. cliffs and high flow creeks).

We walked along the survey line using the image of the route displayed on our GPS units as an aid. We conducted the survey as if it was one continuous line transect, recording one observation per individual bird during a given route-visit, but we restricted the analysis to detections of singing males because females are much more difficult to detect and singing males can be reasonably expected to represent the breeding population. Upon detecting a bird, we took a bearing and waypoint from our location along the line and used a laser range-finder to estimate the distance to the bird. The final plots used for analyses were created by buffering the route by a species-specific truncation distance (Buckland et al., 2001) and clipping the buffered area using the grid cell boundaries (Figure 2). Following the field season, observations were associated with plots by projecting them over the plot polygons in ArcGIS. We established 17 routes and a total of 150 plots, which we surveyed three times on average.

We expected that normalized difference vegetation index (NDVI) would be a good predictor of bird distribution given the relatively distinct and simple structural characteristics of vegetation in our study area including barren slopes, low-shrub tundra, and tall shrub thickets. Using a Landsat 8 image of our study area, we discretized plots using the image's original 30-m resolution and attributed each pixel with its underlying NDVI value. This resulted in an average of 328 pixels per plot. We also extracted the elevation at these grid points from a 5-m digital terrain model. Bird locations were subsequently assigned to the nearest pixel centroid based on the bearing, distance, and waypoint information recorded in the field.

Although data were collected for the entire bird community, we selected the American tree sparrow and savannah sparrow for use as examples because these they had sufficient data for analysis and represented contrasting patterns of distribution which could affect the strength of inference under the TPP model. American tree sparrow is associated with medium-tall shrub thickets and consequently has a clumped distribution, whereas savannah sparrow is associated with open, low-shrub tundra and is more widespread in our study area. For American tree sparrow, we discarded c. 5% of the singing observations at the furthest distances, resulting in a 250 m strip width and a total of 325 detections. For savannah sparrow, discarding c. 5% of the furthest singing observations resulted in plots defined by a 200 m strip width and 1,188 total detections.

For savannah sparrow, we specified the model for the superpopulation point intensity as:

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 * \text{NDVI}_{ij} + \beta_2 * \text{NDVI}_{ij}^2 + \beta_3 * \text{Elev}_{ij} + \beta_4 * \text{Elev}_{ij}^2 + \varepsilon_{\text{subplot}},$$

where $\varepsilon_{\text{subplot}}$ is a mean zero random effect for the 300×300 m quarter sections of each plot. We included the random effects term because unexplained variation in the point intensity function may lead to bias in spatial effect, abundance and variance estimates (Cressie, 1993; Yuan et al., 2016). Due to the comparatively sparse distribution of American tree sparrow, we fit a simplified model with the effect of NDVI, linear and quadratic effects of elevation, and a plot-level random effect.

For savannah sparrow, we described the probability of availability $\phi_{j,k}$ as a linear function of time after 02:00 AM, $\text{time}_{j,k}$, linear and quadratic effects of date, a survey-level random effect $\delta_{j,k}$, and the duration of the plot-visit, $\text{effort}_{j,k}$:

$$\text{logit}(\phi_{j,k}) = \alpha_0 + \alpha_1 * \text{time}_{j,k} + \alpha_2 * \text{date}_{j,k} + \alpha_3 * \text{date}_{j,k}^2 + \alpha_4 * \text{effort}_{j,k} + \delta_{j,k},$$

assuming that as survey effort increases (e.g. time or route length), the population that is exposed to sampling during a survey occasion will approach the 'true' population size (Royle & Dorazio, 2006). We used 02:00 a.m. to calculate the survey timing covariate because there is no sunrise/sunset in June for our study area and this hour corresponds to the initiation of singing for many species. We included the linear and quadratic effects of survey date to account for variation in the probability of availability over a survey period that began prior to the arrival of a subset of individuals and then extended past peak singing (Kéry, Royle, Plattner, & Dorazio, 2009; Schmidt et al., 2013). Lastly, we included a survey-level random effect, $\delta_{j,k}$, to accommodate unexplained variation in the model (Buckland et al., 2016; Oedekoven, Laake, & Skaug, 2015), in part due to the effects of weather. For the smaller American tree sparrow dataset, we fit a reduced model that included only the fixed effects of time and effort.

We also fit a similarly parameterized log-normal TEDS model (with 10 distance classes), except that the spatial effects were reduced to the plot mean and we included an offset (i.e. $\log(\text{plot area}_j)$) in the model for variation in local abundance. We report estimates from both TPP and log-normal TEDS models as posterior means with 95% credible intervals. We also present predictive maps of point intensity from the TPP model.

We fit all models using JAGS (Plummer, 2003). We specified vague priors for all parameters as described in the simulation study section. Summaries of the posterior distribution were calculated from two independent Markov chains run for 100,000 and 75,000 iterations (savannah and American tree sparrow, respectively) with a 25,000 iteration burn-in and thinning every five draws. We used the Gelman-Rubin diagnostic (Brooks & Gelman, 1998) to assess convergence and scaled continuous covariates ($M = 0$, $SD = 1$) to improve convergence properties. See Appendix S3 for JAGS model code.

3 | RESULTS

Based on data generated under the TPP model, the Poisson TEDS model exhibited negative bias in population size estimates and poor credible interval coverage for all parameters (Table 1). Bias in parameter and abundance estimates was relatively low for both the log-normal TEDS model (−3.5 to 5.6%) and the TPP model (−0.1 to 1.9%). In contrast to the log-normal TEDS model, credible interval coverage for all parameters estimated under the TPP model was approximately nominal (Table 1). Under the log-normal TEDS model, coverage for σ was particularly poor (0.73). In addition, the RMSE of the σ estimates (truth = 3) was larger for the log-normal TEDS model (0.30 vs. 0.17). Lastly, the average coefficient of variation (CV) of the total population estimate was 26% smaller for the TPP model compared to the log-normal TEDS model.

For the arctic passerine application, the log-normal TEDS model explained very little of the spatial variation in abundance for both savannah and American tree sparrow, in strong contrast to inferences from the TPP model. Between both species, all but one of the 95%

TABLE 1 Results from 500 simulations for Poisson TEDS, Poisson log-normal TEDS, and TPP models where σ is the scale parameter of the half-normal detection function, ϕ is the probability of availability, β_1 is a spatial effect on the point intensity and M_{total} is the total superpopulation size (a derived parameter) which varied between simulation runs due to randomly drawing the local population sizes

Parameter	Model	True value	M	RMSE	Bias	CI coverage
σ	TEDS	3	3.11	0.31	0.04	0.72
	Log-normal TEDS	3	3.11	0.30	0.04	0.73
	TPP	3	3.06	0.17	0.02	0.95
ϕ	TEDS	0.6	0.66	0.08	0.10	0.70
	Log-normal TEDS	0.6	0.58	0.06	-0.04	0.89
	TPP	0.6	0.60	0.04	0.00	0.96
β_1	TEDS	-	2.03	-	-	-
	Log-normal TEDS	-	2.11	-	-	-
	TPP	2	2.00	0.04	0.00	0.95
M_{total}	TEDS	-	303.84	30.55	-0.08	0.66
	Log-normal TEDS	-	354.74	32.58	0.06	0.93
	TPP	-	337.73	20.29	0.01	0.95

CIs for the spatial effects from the log-normal TEDS model covered 0, whereas none of 95% CIs for the same effects from the TPP model covered 0 (Table 2). In addition, the CVs for the population estimates from the TPP were 32% and 4% smaller for savannah and American tree sparrow, respectively. For both species, population estimates from the TPP model were smaller than those from the log-normal TEDS model (Table 2). Lastly, there was considerable variation in the available population across survey occasions that was apparently well explained by temporal and search intensity covariates (Table 2). That is, the survey-specific expected $\phi_{j,k}$ ranged from 0.20 to 0.89 and 0.17 to 0.95 for savannah and American tree sparrow, respectively.

The fine spatial grain of the TPP model revealed strong and contrasting habitat relationships between American tree sparrow and savannah sparrow (Figure 3). As expected, American tree sparrow exhibited a clumped distribution, corresponding to the non-uniform distribution of their preferred habitat, medium to tall shrub thickets. In contrast, savannah sparrow was more uniformly distributed in a landscape largely comprised of lower vegetative growth forms (Figure 3). The TEDS model generally failed to identify patterns of habitat association in either species (Table 2), and permitted inference about variation in abundance at a much coarser spatial grain (i.e. the plot).

4 | DISCUSSION

Ecological studies are inherently spatial; integrating spatial processes in sampling design, encounter probability and population dynamics (Borchers & Marques, 2017; Royle, Chandler, Sollmann, & Gardner, 2013). However, the sensitivity of studies to spatial processes may be affected by choice of inference method including their corresponding assumptions about the observation process (Nichols et al., 2009). Herein, we used an approach that frames the sampling design and analysis in terms of individual encounters in space rather than aggregating counts at the plot-level (Kéry & Royle, 2016). While distance sampling models that incorporate plot-level spatial effects will often be a good description of variation in abundance (Kéry & Royle, 2016), our application demonstrates that spatial distance sampling models

can provide considerably stronger inference in the presence of within-plot variation in density.

The lower precision (i.e. a 36% larger average CV for the population estimate) of the Poisson log-normal TEDS model relative to the TPP model in our simulation study was expected given that the model for variation in the local population sizes was a reduced information summary of the data generating process and there was greater uncertainty in explaining the remaining variation in abundance as random noise. The negative bias of population size estimates from the Poisson TEDS model lacking an over-dispersion term was likely due to a failure to fit some local population sizes that were larger than would be expected under the estimated Poisson mean. In addition, our simulation study indicated that the improper assumption of a homogeneous point pattern may result in relatively poor estimation of σ .

In the case studies, TEDS and TPP models provided contrasting inference. For both species, the log-normal TEDS model indicated weak relationships between plot-level abundance and habitat, whereas the TPP model indicated strong relationships for those same effects on point intensity, suggesting the existence of relatively strong within-plot variation in density (i.e. inhomogeneous distribution of vegetation). In open landscapes like those in our Alaskan study area, within-plot variation in bird density may be considerable due to the need to establish plots that are sized appropriately relative to a large effective sampling area coupled with variation in vegetation that is compressed along steep elevational gradients. Under these situations, the TPP model would be expected to provide improved inference due to properly specifying an inhomogeneous distribution of individuals within plots and leveraging the spatial information inherent in the location of encounter.

In addition to explicitly addressing spatial processes affecting the distribution of individuals, our approach also accommodates non-Markovian temporary emigration. In single-visit distance sampling surveys, abundance and detection probability are conditional on the probabilities of presence and availability, p_p , and p_a , respectively. However, variation in the temporary emigration composite process, $p_p p_a$, may be confounded to an unknown degree with that present in

TABLE 2 Parameter estimates and 95% credible intervals for Poisson log-normal Temporary emigration distance sampling (TEDS) and Thinned point process (TPP) models. Estimates for α parameters correspond to effects on ϕ and are on the logit-scale. Estimates for β parameters correspond to effects on the point intensity and are on the log-linear scale. M_{total} is the total superpopulation and σ is the scale parameter of a half-normal detection function. Bold numbers indicate estimates with 95% credible intervals that do not include 0

Species	Parameter	Log-normal TEDS		TPP	
		M	95% CI	M	95% CI
American tree sparrow	σ	135.93	(119.75, 157.44)	156.89	(132.54, 192.43)
	α_0	0.22	(-0.48, 1.02)	-0.01	(-0.61, 0.55)
	α_1 *time	0.26	(-0.03, 0.59)	0.21	(-0.03, 0.46)
	α_2 *effort	0.80	(0.41, 1.35)	0.64	(0.34, 1.01)
	β_0	-1.93	(-5.23, 1.32)	-6.24	(-6.65, -5.81)
	β_1 *NDVI	1.38	(0.73, 2.08)	1.93	(1.64, 2.22)
	β_2 *elev	0.12	(-0.17, 0.41)	-0.33	(-0.62, -0.06)
	β_3 *elev ²	-0.88	(-2.70, 0.92)	-0.488	(-0.79, -0.22)
	M_{total}	252.55	(202, 336)	243.826	(199, 321)
Savannah sparrow	σ	113.20	(105.24, 122.24)	116.003	(107.52, 126.09)
	α_0	-0.49	(-1.07, 0.02)	-0.347	(-0.78, 0.07)
	α_1 *time	-0.20	(-0.34, -0.08)	-0.22	(-0.36, -0.09)
	α_2 *date	0.14	(0.00, 0.30)	0.225	(0.07, 0.40)
	α_3 *date ²	0.19	(0.06, 0.33)	0.224	(0.09, 0.37)
	α_4 *effort	0.50	(0.32, 0.73)	0.515	(0.33, 0.73)
	β_0	-2.68	(-5.34, -0.02)	-4.186	(-4.44, -3.90)
	β_1 *NDVI	-0.05	(-0.40, 0.29)	1.193	(1.00, 1.40)
	β_2 *NDVI ²	-0.03	(-0.69, 0.62)	-0.497	(-0.69, -0.31)
	β_3 *elev	0.11	(-0.07, 0.30)	-0.683	(-0.80, -0.57)
	β_4 *elev ²	1.06	(-0.92, 3.05)	0.141	(0.02, 0.26)
	M_{total}	1368.80	(1045, 1947)	1233.76	(1018, 1558)

abundance, negatively biasing abundance estimates (Chandler et al., 2011; Schmidt et al., 2013). Temporary emigration processes are particularly problematic for surveys of species that show large variation in cue production over a survey season and highly mobile species, including: birds (e.g. Chandler et al., 2011; Nichols et al., 2009; Schmidt et al., 2013), herpetofauna (e.g. O'Donnell, Thompson, & Semlitsch, 2015) and insects (e.g. Kéry et al., 2009). For example, in our study, the proportion of American tree and savannah sparrows that were available for sampling on a given visit ranged from c. 0.2 to nearly 1. Using repeat surveys with inference to the superpopulation allowed us to directly estimate $p_p p_a$ and, thereby, make minimal untested assumptions regarding the observation process. Theoretically, the thinned point process model described here could be combined with the distance sampling time-removal model of Amundson, Royle, and Handel (2014) to explicitly estimate all of the primary components of detection. However, for our arctic passerine application, it is unlikely that the combined model could fully separate p_p and p_a . Much of the information about temporary emigration, particularly seasonal variation in song frequency, comes from visits spaced several days apart (rather than the time intervals within a visit) in part due to the Markovian nature of birdsong (Collins, 2004). Consequently, we would expect that the estimated p_p under this

combined model would still be some version of a composite $p_p p_a$ comprising a proportion of p_a that cannot be fully explained by the time-removal model. For our application, we would expect that directly accommodating individual heterogeneity in p_p , in the form of variation in the proximity of home ranges to the survey transect, would be necessary for separate estimation of p_a . This type of inference would be possible within the framework of the spatial count model (Chandler & Royle, 2013) which extends the spatial point process underlying spatial capture-recapture models (Borchers & Efford, 2008; Efford, 2004; Royle & Young, 2008) to unmarked and partially marked populations.

Studies of population dynamics in remote areas face considerable logistical challenges making it critical to develop efficient survey designs for these landscapes. Particularly in rugged terrain, a completely randomized design, as required under the uniformity assumption of conventional distance sampling, will often lead to inefficiency (Marques, Buckland, Borchers, Tosh, & McDonald, 2010). Spatial distance sampling inference permits non-standard designs where the focus can be on maximizing the population that is exposed to sampling in an efficient manner. For our study, this consideration was particularly important given that many arctic passerines are sparsely distributed and the terrain is not amenable to a completely random

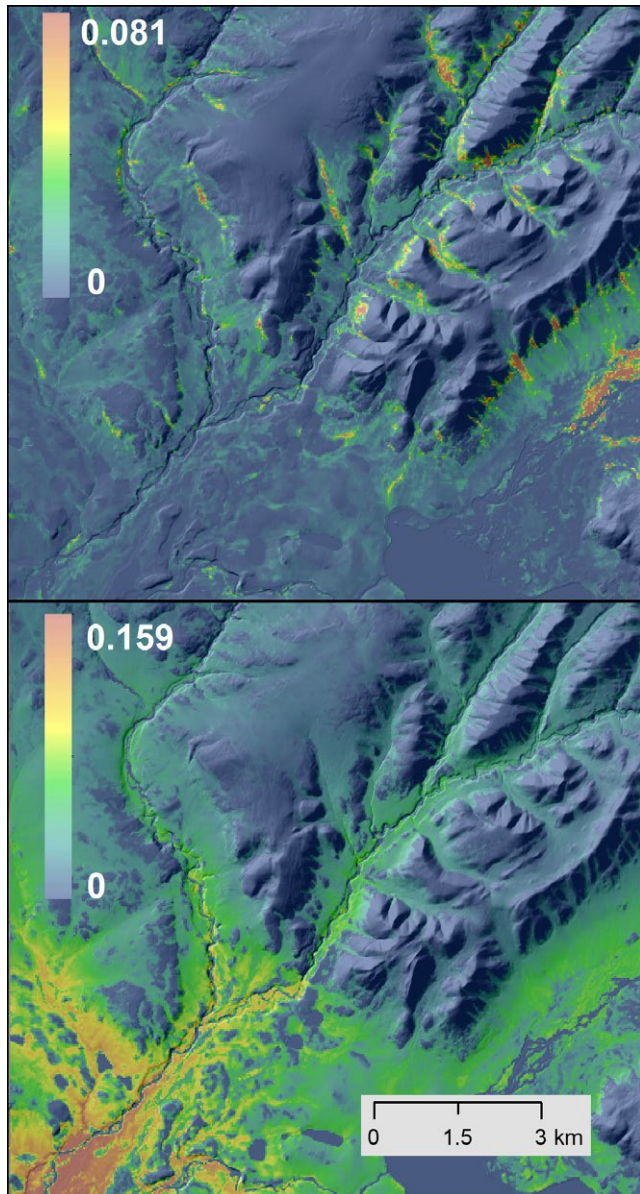


FIGURE 3 Posterior predictions of superpopulation point intensity (abundance per pixel) for American tree sparrow (top) and savannah sparrow (bottom)

and efficient transect configuration. Thus, configuring transects such that we limited our time spent “off-effort” greatly increased the amount of data we were able to collect within the limited breeding season.

Framing inference in terms of the individual encounter in space offers greater sensitivity for understanding spatial processes particularly in the presence of compressed habitat gradients (Royle et al., 2013). However, in many ecological studies, investigators are confronted with large variation in the available population across survey occasions. Thus, we expect that accounting for temporary emigration in spatial distance sampling models may be a critical specification for many species. That is, to fully realize the spatial information inherent in the location of encounter, it may be necessary to directly estimate $p_p p_a$ or p_p . We expect that the methods described here will be particularly

applicable to monitoring of species undergoing concurrent changes in distribution and abundance, particularly the study of populations subject to changes in abundance that are being driven by changes in habitat distribution.

ACKNOWLEDGEMENTS

We thank J.B. Hughey, S.A. Swanson, S.C. Swanson, and E. Burk for conducting surveys. We thank J.A. Royle for initially suggesting a point pattern approach and J.P. Lawler for his support during the development of the landbird monitoring program. We also wish to thank G.A. Breed, A.N. Powell, E.J. Wald, and two anonymous reviewers for valuable comments on an earlier version of this manuscript. Funding for this work was provided through the U.S. National Park Service’s Inventory and Monitoring Program and the Arctic Network. Any reference to trade names or commercial products are for descriptive purposes only and do not imply endorsement by the U.S. government and the research conducted in this study is in compliance with its current laws.

AUTHORS’ CONTRIBUTIONS

J.M. conceived the ideas and designed methodology; J.M. collected and analysed the data; J.M., J.S. and M.L. wrote the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

The data from the arctic passerine application are available at the Dryad Digital Repository <https://doi.org/10.5061/dryad.90kp2> (Mizel, Schmidt, & Lindberg, 2017).

ORCID

Jeremy D. Mizel  <http://orcid.org/0000-0002-9778-7609>

Joshua H. Schmidt  <http://orcid.org/0000-0002-5803-6431>

REFERENCES

- Amundson, C. L., Royle, J. A., & Handel, C. M. (2014). A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. *The Auk*, 131, 476–494. <https://doi.org/10.1642/AUK-14-11.1>
- Borchers, D. L., & Efford, M. G. (2008). Spatially explicit maximum likelihood methods for capture–recapture studies. *Biometrics*, 64, 377–385. <https://doi.org/10.1111/j.1541-0420.2007.00927.x>
- Borchers, D. L., & Marques, T. A. (2017). From distance sampling to spatial capture–recapture. *ASTA Advances in Statistical Analysis*, 101, 475–494.
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, 434–455.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). *Introduction to distance sampling*. Oxford, UK: Oxford University Press.

- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2004). *Advanced distance sampling*. Oxford, UK: Oxford University Press.
- Buckland, S. T., Oedekoven, C. S., & Borchers, D. L. (2016). Model-based distance sampling. *Journal of Agricultural, Biological & Environmental Statistics*, 21, 58–75. <https://doi.org/10.1007/s13253-015-0220-7>
- Chandler, R. B., & Royle, J. A. (2013). Spatially explicit models for inference about density unmarked or partially marked populations. *The Annals of Applied Statistics*, 7, 936–954. <https://doi.org/10.1214/12-AOAS610>
- Chandler, R. B., Royle, J. A., & King, D. I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology*, 92, 1429–1435. <https://doi.org/10.1890/10-2433.1>
- Collins, S. (2004). Vocal fighting and flirting: The functions of birdsong. In P. Marler, & H. Slabbekoorn (Eds.), *Nature's music: The science of birdsong* (pp. 69–72). San Diego, CA: Elsevier Academic Press.
- Cressie, N. A. C. (1993). *Statistics for spatial data: Wiley series in probability and mathematical statistics*. Hoboken, NJ: Wiley.
- Dial, R. J., Smeltz, T. S., Sullivan, P. F., Rinas, C. L., Timm, K., Geck, J. E., Tobin, S. C., ... Berg, E. C. (2015). Shrubline but not treeline advance matches climate velocity in montane ecosystems of south-central Alaska. *Global Change Biology*, 22, 1841–1856.
- Efford, M. G. (2004). Density estimation in live-trapping studies. *Oikos*, 106, 598–610. <https://doi.org/10.1111/j.0030-1299.2004.13043.x>
- Hedley, S. L., & Buckland, S. T. (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, 9, 181–199. <https://doi.org/10.1198/1085711043578>
- Hiby, L., & Krishna, M. B. (2001). Line transect sampling from a curving path. *Biometrics*, 57, 727–731. <https://doi.org/10.1111/j.0006-341X.2001.00727.x>
- Johnson, D. S., Laake, J. L., & Ver Hoef, J. M. (2010). A model-based approach for making ecological inference from distance sampling data. *Biometrics*, 66, 310–318. <https://doi.org/10.1111/j.1541-0420.2009.01265.x>
- Kellner, K. (2016). jagsUI: A wrapper around rjags to streamline JAGS analyses. *R package version*, 1(1).
- Kendall, W. L., Nichols, J. D., & Hines, J. E. (1997). Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology*, 78, 563–578.
- Kéry, M., & Royle, J. A. (2016). *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS, Volume 1: Prelude and static models*. Waltham, MA: Academic Press.
- Kéry, M., Royle, J. A., Plattner, M., & Dorazio, R. M. (2009). Species richness and occupancy estimation in communities subject to temporary emigration. *Ecology*, 90, 1279–1290. <https://doi.org/10.1890/07-1794.1>
- Marques, T. A., Buckland, S. T., Borchers, D. L., Tosh, D., & McDonald, R. A. (2010). Point transect sampling along linear features. *Biometrics*, 66, 1247–1255. <https://doi.org/10.1111/j.1541-0420.2009.01381.x>
- Miller, D. L., Burt, M. L., Rexstad, E. A., & Thomas, L. (2013). Spatial models for distance sampling data: Recent developments and future directions. *Methods in Ecology and Evolution*, 4, 1001–1010. <https://doi.org/10.1111/2041-210X.12105>
- Mizel, J. D., Schmidt, J. H., & Lindberg, M. L. (2017). Data from: Accommodating temporary emigration in spatial distance sampling models. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.90kp2>
- Mizel, J. D., Schmidt, J. H., McIntyre, C. L., & Roland, C. A. (2016). Rapidly shifting elevational distributions of passerine species parallel vegetation change in the subarctic. *Ecosphere*, 7, e01264. <https://doi.org/10.1002/ecs2.1264>
- Nichols, J. D., Thomas, L., & Conn, P. B. (2009). Inferences about landbird abundance from count data: Recent advances and future directions. In D. L. Thompson, E. G. Cooch, & M. J. Conroy (Eds.), *Modeling demographic processes in marked populations* (pp. 201–235). New York, NY: Springer. <https://doi.org/10.1007/978-0-387-78151-8>
- O'Donnell, K. M., Thompson, F. R. III, & Semlitsch, R. D. (2015). Partitioning detectability components in populations subject to within-season temporary emigration using binomial mixture models. *PLoS ONE*, 10, e0117216. <https://doi.org/10.1371/journal.pone.0117216>
- Oedekoven, C. S., Laake, J. L., & Skaug, H. J. (2015). Distance sampling with a random scale detection function. *Environmental and Ecological Statistics*, 22, 725–737. <https://doi.org/10.1007/s10651-015-0316-9>
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing. (DSC 2003). March, 20–22.
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60, 108–115. <https://doi.org/10.1111/j.0006-341X.2004.00142.x>
- Royle, J. A., Chandler, R. B., Sollmann, R., & Gardner, B. (2013). *Spatial capture-recapture*. Waltham, MA: Academic Press.
- Royle, J. A., Dawson, D., & Bates, S. (2004). Modeling abundance effects in distance sampling. *Ecology*, 85, 1591–1597. <https://doi.org/10.1890/03-3127>
- Royle, J. A., & Dorazio, R. M. (2006). Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 249–263. <https://doi.org/10.1198/108571106X129153>
- Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology*. London, UK: Academic Press.
- Royle, J. A., & Young, K. V. (2008). A hierarchical model for spatial capture-recapture data. *Ecology*, 89, 2281–2289. <https://doi.org/10.1890/07-0601.1>
- Schmidt, J. H., McIntyre, C. L., & MacCluskie, M. C. (2013). Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? *Biological Conservation*, 160, 130–139. <https://doi.org/10.1016/j.biocon.2013.01.007>
- Sturm, M., Racine, C., & Tape, K. (2001). Climate change: Increasing shrub abundance in the Arctic. *Nature*, 411, 546–547. <https://doi.org/10.1038/35079180>
- Tape, K., Sturm, M., & Racine, C. (2006). The evidence for shrub expansion in Northern Alaska and the Pan-Arctic. *Global Change Biology*, 12, 686–702. <https://doi.org/10.1111/j.1365-2486.2006.01128.x>
- Yuan, Y., Bachl, F. E., Lindgren, F., Brochers, D. L., Illian, J. B., Buckland, S. T., ... Gerrodette, T. (2016). *Point process models for spatio-temporal distance sampling data*. arXiv preprint arXiv:1604.06013.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Mizel JD, Schmidt JH, Lindberg MS. Accommodating temporary emigration in spatial distance sampling models. *J Appl Ecol*. 2018;55:1456–1464. <https://doi.org/10.1111/1365-2664.13053>