

# Integrated distance sampling models for simple point counts

Kery, Marc; Royle, J. Andrew; Hallman, Tyler; Robinson, W. Douglas ; Strebel, Nicolas; Kellner, Kenneth F.

**Ecology**

*DOI:*
10.48550/arXiv.2211.17229

Accepted/In press: 12/01/2024

Peer reviewed version

*Ecology:* Article

# Integrated distance sampling models for simple point counts

**Short title:** Integrated distance sampling

Marc Kéry[1,*], J. Andrew Royle[2], Tyler Hallman[1,5,6], W. Douglas Robinson[3], Nicolas Strebel[1], Kenneth F. Kellner[4]

[1] Swiss Ornithological Institute, Seerose 1, 6204 Sempach, Switzerland

[2] USGS Eastern Ecological Science Center, Laurel, 20708 Maryland

[3] Oak Creek Lab of Biology, Department of Fisheries, Wildlife, and Conservation Science, Oregon State University, Corvallis, Oregon USA;

[4] Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan USA;

[5] Department of Biology and Chemistry, Queens University of Charlotte, Charlotte, North Carolina USA;

[6] School of Environmental and Natural Sciences, Bangor University, Bangor, LL57 2UW, UK

* corresponding author: marc.kery@vogelwarte.ch

**Open Research**

Data and code used in the case study, along with R and JAGS code for conducting simulation studies 1–4 described in the manuscript, will be posted on Zenodo upon acceptance; they are currently on GitHub: `https://github.com/kenkellner/IDS`. The dev version of R package `unmarked` which contains the new IDS model fitting function `IDS()` is available on GitHub and can be installed from R by submitting the following command:

`remotes::install_github("kenkellner/unmarked", ref="IDS")`.

For the case study, eBird data was downloaded as described in the manuscript and in Appendix S2: Section S2.

**Key words:** abundance; availability probability; biodiversity monitoring; citizen-science; detection/nondetection data; distance sampling; integrated model; perceptibility; point count data

## Abstract

Point counts (PCs) are widely used in biodiversity surveys, but despite numerous advantages, simple PCs suffer from several problems: detectability, and therefore abundance, is unknown; systematic spatiotemporal variation in detectability yields biased inferences, and unknown survey area prevents formal density estimation and scaling-up to the landscape level. We introduce integrated distance sampling (IDS) models that combine distance sampling (DS) with simple PC or detection/nondetection (DND) data to capitalize on the strengths and mitigate the weaknesses of each data type. *Key to IDS models is the view of simple PC and DND data as aggregations of latent DS surveys that observe the same underlying density process.* This enables estimation of separate detection functions, along with distinct covariate effects, for all data types. Additional information from repeat or time-removal surveys, or variable survey duration, enables separate estimation of the availability and perceptibility components of detectability with DS and PC data. IDS models reconcile spatial and temporal mismatches among data sets and solve the above-mentioned problems of simple PC and DND data. To fit IDS models, we provide JAGS code and the new `IDS()` function in R package `unmarked`. Extant citizen-science data generally lack the information necessary to adjustment for detection biases, but IDS models address this shortcoming, thus greatly extending the utility and reach of these data. In addition, they enable formal density estimation in hybrid designs, which efficiently combine distance sampling with distance-free, point-based PC or DND surveys. We believe that IDS models have considerable scope in ecology, management, and monitoring.


## Introduction

Point count methods are among the most widely used and longest-standing protocols in wildlife surveys worldwide (Rosenstock et al. 2002; Darras et al. 2021). Simple point counts (PC) are brief surveys in which a stationary observer counts all individuals of some species (single species to entire communities) detected either without distance constraints or within a predefined distance

from the observer. Point count methods are logistically uncomplicated and are ubiquitous in biodiversity surveys worldwide, e.g., in the North American Breeding Bird Survey/BBS (Sauer et al. 2017), and many European national BBS or bird atlas schemes (Balmer et al. 2013). In addition, a wide range of what in essence can be conceptualized as point count methods, albeit varying from highly standardized to essentially design-free, is at the core of rapidly growing citizen-science projects such as eBird (Sullivan et al. 2009).

Despite the prevalence of simple point counts, their simplicity is not without drawbacks; e.g., it is not possible to estimate true abundance or occupancy if visits to points are unreplicated (Stoudt et al. 2023). In addition, PC data are non-spatial in the sense that the area from which the detected animals are drawn is usually unknown. This prevents spatial extrapolation for rigorous estimation of regional population sizes. Similarly, integrated analysis of data from different schemes is hampered due to commonly occurring spatial mismatches. Finally, variable survey duration is very common and creates a temporal mismatch in the data thus different data points do not correspond to the same survey effort (Pacifici et al. 2019). Both greatly complicate joint analyses from multiple survey schemes that use point count methods.

In planned surveys, additional information is often collected during PC surveys that permits estimation of detection probability (Nichols et al. 2009). Such extra information includes replicated counts (Royle 2004), double-observer surveys (Nichols et al. 2000), removal counts (Wyatt 2002; Dorazio et al. 2005), distance information (Marques et al. 2007; Buckland et al. 2015), or locational information from recognizable individuals which enables the fitting of spatial capture-recapture (SCR) models (Borchers & Efford 2008; Royle et al. 2014). These survey protocols permit estimation of abundance, and thus assessment of status and trends free from any bias produced by imperfect detection and by unmodelled temporal or spatial patterns in detectability (Kéry & Royle 2016, 2021). While these methods produce detectability-adjusted indices of abundance, only estimates from SCR and DS are area-explicit.

Even if DS or SCR data are available, it is not clear at present how they should be used alongside, or combined with, data from simple PCs available from national BBS or bird atlas schemes. For instance, a major challenge in the joint modeling of such data types is how to address spatial or temporal mismatches (Pacifici et al. 2019); these arise when effective sampling areas are unknown and vary, and when survey durations differ (Solymos et al. 2013). Thus, it would be desirable to have formal methods for combining these different data types that build on their complementary strengths, e.g., detectability estimation, trained observers *vs.* large sample size, geographic breadth, etc.

Here, we introduce integrated distance sampling (IDS) models that permit the combination of data from DS, simple PC, and detection/nondetection (DND) surveys. Our integrated model is based on an underlying hierarchical distance sampling model (Royle et al. 2004; Kéry & Royle 2016: chapter 8) for all three data types. We conceptualize data from all three survey methods as the outcome from a (possibly latent) distance sampling protocol, i.e., where detection probability is assumed to be a function of distance from the observer. This enables us to estimate separate detection functions for each data set, which automatically reconciles any spatial mismatch among the data types and surveys. Temporal mismatches, i.e., variable survey duration, in PC data can be addressed by including an availability process in the model, which is informed by extra data such as variable survey duration, or by multi-observer, replicate or time-removal surveys (Borchers et al. 1998, Diefenbach et al. 2007, Solymos et al. 2013; Amundson et al. 2014). These extra data allow for separation of the availability and perceptibility components of detection probability (Marsh & Sinclair 1989; Nichols et al. 2009, Hostetter et al. 2019, Péron & Garel 2019).

Key to our IDS models is the view of PC and DND data as aggregations, or summaries, of latent DS survey data, with identical density, and possibly availability, processes as regular DS surveys. Hence, we view PC and DND data types as DS counts where distance information is unavailable. As we will show, the key assumption of a shared density and availability process permits estimation of separate detection functions, along with different parameters linking these

functions to covariates, for all three data types. Estimation of separate detection functions, when needed, can accommodate any systematic differences between survey schemes. Combining PC and DND data with the more information-rich DS data enables estimation of detection probability and makes the resulting abundance estimates area-explicit: effective survey areas for PC and DND surveys become estimable, and population density is estimated with improved precision. Thus, IDS models can reconcile all discrepancies, including spatial and temporal mismatches, among these extremely widespread data types.

In this article, we begin by formally describing IDS models. We then use simulation to demonstrate identifiability of our model when separate detection functions are estimated for each data type, including separate parameters for detection function covariates. Next, we explore the effects of adding variable amounts of the more information-rich but "expensive" DS surveys to a larger sample of the less information-rich but "cheaper" PC data. Following that, for a model combining DS and PC data, we demonstrate the identifiability of availability in addition to perceptibility, provided that surveys vary in duration; this is one of the types of extra information which enable availability to be estimated separately from perceptibility (e.g., Solymos et al. 2013). Finally, we showcase IDS models with the Oregon 2020 Project (Robinson et al. 2020) as a case study. As part of the case study, we demonstrate the ability of IDS models to allow for different magnitudes of heterogeneity in the detection functions estimated for different portions of the data (Oedekoven et al. 2015). Such accommodation of intricate, survey-specific features of the observation process may be particularly important when reconciling data from heterogeneous survey protocols in a single integrated model.

We have implemented a range of IDS models in the new fitting function `IDS()` in the R package `unmarked` (Fiske & Chandler 2011, Kellner et al. 2023), to permit user-friendly model fitting by maximum likelihood, and we provide BUGS code for Bayesian inference using JAGS (Plummer 2003). We believe that IDS models have a large scope of application for exploiting PC

and DND data in a more rigorous and synthetic manner, and to obtain less biased and larger-scale inferences about abundance and density, particularly for large citizen-science data sets.

## Integrated Distance Sampling (IDS) models

We develop joint likelihoods, i.e., integrated models (Besbeas et al. 2002; Miller et al. 2019; Kéry & Royle 2021: chapter 10; Schaub & Kéry 2022), for the following data types, which we assume to observe the same density and availability processes. We note that $i$ indexes different sites across data types. Our current models assume closure and the absence of any temporal replicates at a site, but relaxation of both assumptions will be the subject of future work.

(1) Distance-sampling (DS) data $y_{i,j}^{ds}$, possibly with truncation distance $b_i^{ds}$ and survey duration $t_i^{ds}$, where *j* indexes *J* distance classes, and where $y_{i,\cdot}^{ds} = \sum_{j=1}^{J} y_{i,j}^{ds}$ denotes the total count per site.

(2) Simple point counts (PC) $y_i^{pc}$ with duration $t_i^{pc}$, with or without a truncation distance $b_i^{pc}$, as produced by many national BBS or bird atlas schemes.

(3) Detection/nondetection (DND) data $y_i^{dnd}$, indicating the observed presence or absence of a species during a point-location survey of duration $t_i^{dnd}$ out to an optional truncation distance $b_i^{dnd}$, as they are similarly produced by countless biological surveys.

For joint inference about density, first, for the DS data we adopt a hierarchical distance sampling (HDS) model (Royle et al. 2004) represented by $N_i \sim Poisson(A_i \lambda_i)$ and $y_{i,\cdot} \sim Binomial(N_i, \theta_i p_i^{ds})$. Completing the HDS model, the site-specific vector of distance-class counts has a multinomial distribution with cell probabilities computed by integrating the distance function over the prescribed intervals (see Kéry and Royle 2021). $N_i$ and $y_{i,\cdot}$ are, respectively, the latent abundance and observed total count at site $i$, with survey area $A_i$ and density $\lambda_i$, while availability ($\theta$) and perceptibility ($p^{ds}$) are the two components of detection probability (Marsh & Sinclair 1989; Nichols et al. 2009). Perceptibility will primarily be a function of distance and is

estimated from distance data by integrating out to distance $b_i$ a suitable detection function such as a half-normal with scale parameter $\sigma$. Truncation distance $b_i$ defines the survey area, which for a point count survey with perfect detection is $A_i = \pi b_i^2$. This is a key advantage of DS methods: that it associates $N_i$ with a well-defined area. It makes abundance estimates in a DS protocol area-explicit, in contrast to most abundance estimation protocols other than SCR (Borchers & Efford 2008; Royle et al. 2014). For songbirds, the availability probability $\theta$ will be mainly a function of singing rates (Solymos et al. 2013), which cannot be estimated from distance data alone. Hence, conventional distance sampling requires either the assumption of perfect detection at distance 0 or else acceptance that inferences will be restricted to the available part of a population only (Buckland et al. 2015). However, availability becomes estimable in a DS model if certain extra information is collected, e.g., from multiple observers (Borchers et al. 1998), replicated surveys (Chandler et al. 2011), time-removal (Farnsworth et al. 2002), or from variable survey duration, as we will show.

Second, for the PC data we adopt a variant of the binomial *N*-mixture model (Royle 2004), represented by $N_i \sim Poisson(A_i^{pc}\lambda_i)$ and $y_i \sim Binomial(N_i, \theta_i \bar{p}_i^{pc})$. Simple PCs are neither area-explicit nor can detection probability be estimated without temporal replication (Stoudt et al. 2023). This precludes estimation of survey area $A^{PC}$, availability $\theta^{pc}$, and perceptibility $\bar{p}^{pc}$. However, we show how use of PC data alongside regular DS data in an IDS model renders estimable both $A^{PC}$ and $\bar{p}^{pc}$, again via the estimation of the parameters of a suitable detection function. Conceptualizing simple PC data as the outcome from latent DS surveys lets us estimate separate detection functions, along with distinct effects of covariates, for both DS and PC data when they are used as part of an IDS model. Integration of these detection functions over unlimited distance or out to some chosen truncation distance yields the average detection probability $\bar{p}_i^{pc}$ for a PC survey at site $i$ and moreover defines survey area $A^{PC}$. This lets PC surveys contribute information towards estimation of density $\lambda$. This model represents a complete, model-based reconciliation of the spatial mismatch between DS and PC data. In addition, variable survey duration $t^{pc}$ or other extra

information mentioned above may render estimable availability $\theta$ and thus additionally reconciles temporal mismatches among DS and PC surveys as well.

Third, for DND data we adopt a variant of the Royle-Nichols (2003) model. Here, the observed DND data are assumed to follow a Bernoulli distribution with a success probability that depends on both local abundance and parameters of the detection process: $y_i \sim Bernoulli(1-(1-\bar{p}_i^{dnd})^{N_i})$, where $y_i$ denotes a binary DND datum at site $i$ and the other quantities are analogous to the above definitions. As for the PC data, single-visit DND data without any extra information won't normally permit parameter estimation under this model, but we will show how using DND data alongside DS data as part of an IDS model will render identifiable survey area $A_i^{dnd}$ and detection probability $\bar{p}_i^{dnd}$. As for an IDS model with PC data the observation model for DND data can be adjusted for imperfect availability by assuming for the DND datum at each site $i$ $y_i \mid p_1, p_2, ..., p_{N_i} \sim Bernoulli(1-E_p(\prod_j^{N_i}(1-\theta_i^{dnd} p_j)))$, where $E_p$ denotes the expectation and $j$ is an index for the $j=1...N_i$ individuals present. When detection of individuals is independent, this simplifies to $y_i \sim Bernoulli(1-(1-\theta_i^{dnd}\bar{p}_i^{dnd})^{N_i})$. However, we have found availability estimates in a model with DND data to be extremely variable to the extent of being useless (unpublished analyses). This needs further study, but for now we include in the IDS models in our paper either DND data or estimation of availability, but currently not both at the same time. Likewise, the `unmarked` function `IDS()` does not allow estimation of availability in an IDS model that includes DND data.

Our current IDS models always require that some DS data are available and they assume population closure and that all data types observe identical abundance and availability processes. Hence, abundance and, if modelled explicitly (for DS and PC data), availability parameters are shared in a joint likelihood, while detection parameters can be either shared or made specific to each data type. We will show that this enables IDS models to obtain separate intercept and slope estimates in the detection function, and therefore of survey area $A$, density $\lambda$ and detection

probability $p$, from unreplicated, simple PC or DND data, when these are used as part of an IDS model. If PC data are the result of surveys with variable duration, an availability process may also be added in the IDS model. For songbirds, Solymos et al. (2013) express availability as a function of singing (or more generally, activity) rate $\phi$ and of survey duration $t$ as $\theta_i = 1 - \exp(-t_i\phi_i)$. We will show how we can also estimate availability in an IDS model combining DS and PC data, provided that survey duration is variable and sample size sufficiently large. We note that we envision a "hiding behaviour" mechanism underlying imperfect availability (Kéry & Royle 2011: section 2.4).

To summarize, for regular DS data we specify likelihood $L^{ds}$ (Royle et al. 2004), for PC data $L^{pc}$ (Royle 2004), and for DND data $L^{dnd}$ (Royle & Nichols 2003). Importantly, for both PC and DND data, we assume a latent DS observation process protocol and estimate detection probability $p$ by integration of a detection function with parameters that become estimable in an IDS model. Under independence among data sets, i.e., when at most a negligible portion of sites appears in more than one data set, we define the following joint likelihoods for three variants of an IDS model: $L^{IDS1} = L^{ds} \times L^{pc}$ (which we call model IDS1) and $L^{IDS2} = L^{ds} \times L^{dnd}$ (model IDS2) for the combinations of DS with PC or DND data, and $L^{IDS3} = L^{ds} \times L^{pc} \times L^{dnd}$ (model IDS3) for the full three-way combination. These likelihoods can be maximized numerically to obtain MLEs, or we can place priors on their parameters and use MCMC methods to obtain Bayesian posterior inferences. See Appendix S1 for a conceptual outline of IDS models and of how they conceptualize PC and DND data as the outcome of a latent distance sampling observation process.

## Tests and demonstrations of IDS models with simulated and real data

***Simulation 1: Identifiability of separate observation process parameters in IDS1 and IDS2***

To demonstrate identifiability of the IDS models, we analyzed simulated data sets and estimated parameters for separate detection functions in an IDS model with either DS + PC data or DS +

DND data, i.e., in the IDS1 and IDS2 case. We used function `simHDS()` in the R package `AHMbook` to simulate two data sets with DS data from 250 sites, and PC or DND data from another 1000 sites. To obtain PC data, we first generated DS data, and then discarded all distance information, just retaining one count per site, and to produce DND data we additionally quantized the resulting counts. Mean density was kept constant at 1, following our assumption of a shared density process. The scale parameter $\sigma$ in the half-normal detection function was set at 100 m for the DS data and was varied randomly between 10 and 130 m for the PC and DND data sets. Thus, the key criterion for identifiability of our models was how well estimates of $\sigma$ matched their true values in the data simulation. In the submodel for the DS data sets, we chose a truncation distance of 200 m. In this simulation we aimed to establish the identifiability of the new models in their simplest form only. That is, we implicitly assumed availability to be 1 and did not use any covariates in either density or detection. We used JAGS (Plummer 2003) to fit IDS1 or IDS2 to 1000 data sets each.

In Simulation 1B (Appendix S2: Section S1) we extended our investigations of parameter identifiability and estimator performance with model IDS1. We varied all of the following four settings independently according to a response-surface design: average density, detection function scale for both DS and PC data, and the DS truncation distance. We again used JAGS for model fitting.

***Simulation 2: Identifiability with distinct covariate effects in the observation model***

We conducted two sets of simulations to answer the following related questions: (i) Does the IDS model allow DS and PC detection to have different covariate relationships in the detection function? (ii) Are relationships still identifiable if the same covariates are related to both detection and density? We answered these questions by simulating data sets with DS and PC data from 200 and 1000 sites, respectively. Density was governed by an intercept of 1 on the natural scale and an effect of 1 of one covariate ("habitat"). The half-normal detection function $\sigma$ had an intercept of 100 and 150 m on the natural scale for DS and PC data, respectively. In a first analysis we used

`simHDS()` to simulate 1000 data sets with these specifications, and where the half-normal detection function $\sigma$, on the log-scale, was affected by another covariate "wind" by independently drawing two random numbers from a Uniform(-0.5, 0.5) distribution, one for the DS data and the other for the PC data. In the second analysis, we used a modified version of function `simHDS()` to simulate another 1000 data sets with the same specifications as above, except that now we generated log-scale effects of the *same* covariate as for density (i.e., "habitat") by independently drawing two U(-0.5, 0.5) random numbers for the DS and PC data sets as their coefficients. We used the new `IDS()` function in R package `unmarked` to fit the data-generating model. We discarded numerical failures, which we conservatively identified by standard errors that were either NA or >5, or by MLEs that were >10 times their true values.

***Simulation 3: How many DS sites are required to obtain adequate estimates of density?***

We simulated 1000 data sets with PC data from 200 sites, to which we added DS data from 1–100 sites in six mixing ratios. Density was governed by an intercept of 1 on the natural scale, with one habitat covariate with coefficient 1. The detection function $\sigma$ was 70 m in the PC and 100 m in the DS data, and we chose a truncation distance of 200 m in the latter. We generated a total of 6000 data sets (1000 for each level of the mixing ratio factor) and fit the IDS1 model using function `IDS()`, discarding numerical failures based on the same criteria as in Simulation 2.

***Simulation 4: How well can availability be estimated in an IDS model?***

We simulated 1000 data sets that resembled our case study below: each had DS data from 3000 sites, and PC data from either 1000, 3000, or 6000 sites. DS survey duration was kept constant at 5 min, but was varied between 3 and 30 min in PC surveys, with a strong right skew, as found in the case study data. Density was governed by an intercept of 1 on the natural scale and with a habitat covariate with coefficient 1, detection function $\sigma$ was 70 m in the PC and 100 m in the DS data, with a truncation distance of 200 m in the latter. The average singing rates per site varied between 0.1 and 2, corresponding to a probability of 0.1–0.86 to sing at least once over a 5 min interval. We fit IDS1 using the `IDS()` function and discarded numerical failures as in Simulation 2.

286 ***Case study: American Robins in the Oregon 2020 Project***

287 We used the IDS1 model to estimate population density of American Robin (*Turdus migratorius*) in
288 Benton and Polk counties, Oregon. The 3680 $km^2$ area in Western Oregon is bounded on the east by
289 the Willamette River and its floodplain, while the western portions include the Coast Range
290 mountains. Silviculture of coniferous forest is the dominant land use in the mountains. Nearly every
291 square kilometer contains a narrow, lightly travelled road for timber harvest, which allowed access
292 for bird surveys. The eastern floodplain sections contain a mix of agricultural uses, mostly festucoid
293 grass seed fields and orchards, and suburban development.

294 DS surveys were conducted every 0.8 km along accessible roads throughout the study area,
295 and every 200-m in an off-road grid placed over the William L. Finley National Wildlife Refuge,
296 producing a total of 2,912 sites sampled and 2020 American Robins detected (Robinson et al.
297 2020). DS surveys were conducted during the breeding season (April 30–July 11) from 2011 to
298 2013 by WDR. Each survey followed the Oregon 2020 protocol (Robinson et al. 2020), which used
299 5-minute stationary counts initiated between 30 min before sunrise and noon on days with no or
300 little rain. All birds detected by sight or sound were recorded with an estimated distance from the
301 observer (verified with a range finder when possible) following standard distance sampling
302 protocols (Buckland et al. 2015).

303 We combined DS surveys with opportunistically-gathered citizen-science PC data from the
304 eBird database (Sullivan et al. 2009), using checklists from 2011-2017 in Benton and Polk counties.
305 After stringent filtering (see Appendix S2: Section S2) and geographic subsampling, 1060 PCs with
306 819 detections of American Robins were included. We strictly filtered data to include only
307 complete checklists using stationary protocols and personal locations, conducted during the
308 breeding season. We further filtered data to include only checklists with durations between 3–30
309 minutes that were conducted between sunrise and seven hours after sunrise. Finally, we applied
310 geographic subsampling to reduce the effects of highly popular sites by overlaying a 200m grid

311 over the study area and randomly selecting only a single checklist from each grid cell. See
312 Appendix S2: Section S2 for eBird query details.
313 For DS data, we selected a truncation distance $b^{ds}$ of 300 meters. We binned the distance data
314 into 50 m distance classes. For the analysis of PC data, an upper distance limit $b^{pc}$ of 500 meters was
315 adopted, assuming that observers do not detect individuals further away than that (the 0.99 quantile
316 in the Oregon 2020 database (Robinson et al. 2020) was 400 m). For both data types, we assumed
317 identical parameters for annual density and availability. We modelled density $\lambda$ with a random
318 intercept for year, and with quadratic terms for elevation and percentage of canopy cover in a 315 m
319 radius around the observer location. This radius was selected as it was previously found to be the
320 most predictive of abundance for this species of the radii considered (Hallman & Robinson 2020).
321 For availability, we adopted the model of Solymos et al. (2013) linking availability probability with
322 activity rate $\phi$ according to a Poisson point process in time, and used linear and quadratic terms for
323 day of the year and minutes since dawn on the log activity rate.
324 We hypothesized that the observation process in the designed DS surveys in the Oregon
325 2020 project might differ from point count surveys recorded in eBird, even after very stringent
326 filtering, as the distance sampling surveys conducted by a professional ornithologist might have a
327 higher detection probability than eBird surveys conducted by citizen scientists with variable training
328 and experience. Therefore, we allowed for different detection functions for the DS and the PC
329 portions in our analysis by fitting separate intercepts in the half-normal detection scale $\sigma$.
330 Moreover, to accommodate possibly different levels of detection heterogeneity among sites, we
331 specified site-specific random effects in $\sigma$ and allowed for a different variance in the DS and PC
332 portions of the data (Oedekoven et al. 2015). In addition, we modeled $\sigma$ using the percentage of
333 urban area and percentage of canopy cover, both in a 165 m radius around the observer location;
334 these slope parameters were shared between DS and PC data. We computed the canopy cover
335 covariate for a smaller radius in the detection function, as the distance that an observer can detect is
336 impacted more heavily by nearby environmental conditions. Elevation, urban land cover, and

canopy cover were obtained from the Oregon Spatial Data Library (Oregon Spatial Data Library 2017), the USGS's National Gap Analysis Project (United States Geological Survey 2011), and Landscape Ecology, Modeling, Mapping and Analysis's gradient nearest neighbor structure maps (LEMMA 2014), respectively.

We processed data in R (R Core Team 2019) and fitted the model in JAGS, using the R package `jagsUI` (Kellner 2016). For all parameters, we chose vague priors; see BUGS model on Zenodo for details (Kéry et al. 2023). We assessed the model goodness-of-fit for both data portions separately using posterior predictive checks (Conn et al. 2018) with a Freeman-Tukey discrepancy measure computed for observed and expected counts for the DS and PC data (Kéry & Royle 2016). This suggested an adequate fit of the model overall: Bayesian p-values for the DS part of the model revealed slight underdispersion, while the PC part of the data indicated good model fit (Appendix S2: Table S6). We obtained posterior predictive distributions of abundance and predicted density, based on elevation and canopy cover, for each of the 3874 1-km$^2$ grid cells in Benton and Polk counties, resulting in an abundance-based species distribution map of American Robin. We also fit a simpler variant of the model using the `IDS()` function in `unmarked` (Kellner et al. 2023) to illustrate both Bayesian and maximum likelihood inference. Code and data to replicate the case study can be found on Zenodo (Kéry et al. 2023).

## Results

***Simulation 1: Identifiability of separate observation process parameters in IDS1 and IDS2***

In an IDS model, separate detection functions were clearly estimable under both IDS1 (combining DS and PC data) and IDS2 (combining DS and DND data); see Fig. 1 and Appendix S2: Table S1. There was no indication of bias in either model: % relative bias was <<1% for all sigma's and <2% for the abundance estimates at sites with $N$>0. Credible interval (CRI) coverage was close to the nominal level of 95% for all estimators. Not surprisingly, precision was slightly lower in model IDS2 than in IDS1 (see middle of Fig. 1). In addition, simulation 1B confirmed the excellent

363 frequentist operating characteristics of the estimators in IDS models under an even wider range of
364 conditions (Appendix S2: Section S1, Appendix S2: Table S2).

365 ***Simulation 2: Identifiability with distinct covariate effects in the observation model***

366 In the first set of simulations, where two distinct covariates affected density and the detection
367 function, and where the effects on the latter were distinct for the DS and PC portions of the data, we
368 discarded 23 sets of estimates as numerical failures. The remaining 977 sets of estimates indicated
369 that this model was identifiable and produced little or no bias (Fig. 2, left; Appendix S2: Table S3
370 left). In the second set of simulations, where the same covariate independently affected density and
371 the two detection functions, we discarded 66 invalid sets of estimates. The remainder again showed
372 this model to be identifiable (Fig. 2, right; Appendix S2: Table S3 right).

373 ***Simulation 3: How many DS sites are required to obtain adequate estimates of density?***

374 In our simple simulation, the IDS model showed excellent performance with as few as 20 DS sites
375 (Fig. 2), with relative bias <1% for all estimators and CI coverage at or near nominal levels
376 (Appendix S2: Table S4). However, the number of numerical failures increased greatly when
377 decreasing numbers of DS data were added in the integrated model; from only 2 out of 1000 when
378 100 DS sites added, to 85 with 20 DS sites, and to 490 out of 1000 when 1 DS site was added.

379 ***Simulation 4: How well can availability be estimated in an IDS model?***

380 Sampling distributions of density estimators were all concentrated around the true value. There
381 were long right tails, but these became more symmetrical with larger sample sizes. Singing rate ( $\phi$
382 ) estimators were precise up to values of about 0.8, 1.3 and 1.4, respectively, for 1000, 3000 and
383 6000 PC sites, but became very imprecise for greater values of the singing rate. Presumably, this
384 was because overall availability reached an asymptote close to 1 when singing rates were very high,
385 making precise estimation of $\phi$ difficult (Fig. 3). Overall, there was a slight positive bias in both
386 density and singing rates, but it declined from 14 to 10% with 1000 PC sites to 3000 and 2% with
387 6000 sites, while CI coverage was always at nominal levels (Appendix S2: Table S5). Relative bias
388 of the detection function scale $\sigma$ for both data types was always less than 1%.

389 ***Case study: American Robins in Oregon***

390 Over all surveys considered, mean survey date was June 7, and time since dawn ranged 17–519 min
391 (mean 229). At mean date and time since dawn, availability within a one minute survey was
392 estimated at 0.295 (95% CRI 0.133–0.795; Appendix S2: Table S6). Estimated availability peaked
393 soon after dawn, decreased during the next five hours, then increased again, and tended to increase
394 slightly throughout the season. Density was estimated to be highest on plots with a canopy cover of
395 ~40% and to decrease with elevation. Median estimates for density varied between 2 and 59
396 individuals per $km^2$. Maxima were found in the foothills where open woodlands transition from the
397 floodplain agricultural zones into the denser forests at higher elevations, while minima were found
398 in the most intensively harvested woodlands (Fig. 5). Over the entire study area, we estimated a
399 population size of 92,439 American Robins (95% CRI 57,322–142,656). Interestingly, on average
400 the estimated detection function scale ($\sigma$) was not different between the DS and PC portions of the
401 data (parameter 'mean.sigma' in Appendix S2: Table S6). However, there was greater variability in
402 the detection function $\sigma$ among surveys on eBird than for regular DS surveys conducted within the
403 Oregon 2020 project (parameter 'sd.eps').

404

405 ## Discussion

406 We discovered how simple point count (PC) or detection/nondetection (DND) data can be formally
407 integrated in a model together with distance sampling (DS) data, to estimate separate parameters of
408 an underlying latent DS observation process in every data type. This allows estimation of a full
409 complement of detection probability parameters for all three data types. Moreover, integrating DS
410 data makes abundance estimates from PC and DND data area-explicit. Thereby, IDS models
411 achieve a formal spatial calibration of PC and point-indexed DND data, as well as a reconciliation
412 of spatial mismatches between all three data types. Thus, IDS models solve two major problems
413 that plague simple point count surveys producing PC or DND data: that detection probability and
414 effective survey areas are both unknown. The key assumption of our IDS model is a shared density

process: that either density is identical among all sample locations, or that density differences can be explained by identical covariate regressions for all data types in the integrated model. These assumptions should be reasonable when all data types are collected randomly in the same general area, and they may also hold when data sets are from disjoint regions, provided some form of random spatial sample is achieved. However, as in perhaps all cases where different data sets are combined in a single analysis, this kind of exchangeability is a judgment call on the part of the analyst. For instance, joining data sets from two spatially-biased samples (e.g., road-side and river-side counts) would not be such a good idea.

We believe IDS models have a large scope of application and can facilitate use of the large amounts of currently available PC data, such as the North American BBS (Sauer et al. 2017) in more formal analyses of abundance that account for imperfect detection. They may also be applied for carefully quality-controlled eBird data (Sullivan et al. 2009), as illustrated in our case study. In the context of our simulation with a Null model without covariates, we have shown that only a relatively small amount of regular DS data was required to supplement simple PC data when used in an IDS model. Our findings agree well with related work with other types of integrated models that demonstrate the benefits of combining even small amounts of data with a higher information content with less informative, but cheaper data (Dorazio 2014, Zipkin et al. 2017, Doser et al. 2021). This would suggest that the scope of inference of point count surveys may be substantially increased by adding even a relatively small number of sites where the additional distance information is collected on purpose. Although – it may well be that in 'real-life', with messy data and consequently with more complex models, (much) more of the information-rich DS data will be needed. This might be addressed with more customized simulations.

In our case study we found that the perceptibility part of detection probability was not different on average between the DS data contributed by the Oregon 2020 project and the PC data obtained from eBird: the intercepts of the detection function scale parameter $\sigma$ were no different between these two portions of the data. However, allowing for random variation of the detection

function parameter $\sigma$ among surveys (i.e., among sites) and for possibly different magnitudes of that variation between the two types of data revealed greater heterogeneity among surveys in the eBird data base than among surveys in the Oregon 2020 project. This makes intuitive sense, since the consistency between surveys must have been higher in the Oregon 2020 project than in eBird: most of our DS data were produced by just one person (WDR), while the eBird data were contributed by many different observers. In addition, our case study emphasizes how careful modeling of patterns in the detection function of an IDS model can help to make data from different protocols more 'alike', by explicitly allowing for their differences in terms of the observation processes that produced them. This is a great strength of IDS models and of parametric statistical inference in general.

Many survey data typically have large variation in duration (Solymos et al. 2013) and thus there is also a need for temporal mismatch among datasets to be addressed. We conceive of this as an availability process (Kendall et al. 1997, Diefenbach et al. 2007), where over time an activity such as singing puts individuals at increasing risk of being detected. Hence, survey duration is naturally informative about availability. However, this part of our model presents more challenges. With the current formulation of our IDS model, we could only estimate availability when combining DS with PC data, but not when DND data were part of the analysis (unpublished analyses), and even then only with large sample sizes. In addition, population closure is required and hence surveys should probably not be very long in duration. Moreover, this part of an IDS model has the form of a single-visit occupancy or N-mixture model (Lele et al. 2012), where estimability hinges upon a continuous, "private" covariate that affects detection, and in our case, availability. Such models are identifiable (Dorazio 2012), but they rely strongly on parametric assumptions and may lack robustness to violations of those assumptions (Stoudt et al. 2023). Our simulation 4 and the case study both showed availability to be identifiable in an IDS1 model, when extra information about the availability process (in our case, variable survey duration) was included. However, our study species was chosen specifically to be fairly common. In rarer species and

consequently smaller sample sizes, there may well be challenges when attempting to estimate availability parameters in an IDS model. When information to estimate availability is too sparse, estimates may tend towards the boundary of 1, which will cause underestimation of density. Finally, we point out that for survey duration there is a tradeoff, since variability in survey duration in the PC data is needed as the source of information about availability. On the other hand, too long survey durations may lead to violations of the closure assumption and (presumably) to an overestimate of density. This is something to keep in mind when planning to apply IDS modeling.

Therefore, IDS models that estimate availability must be developed and applied with care. Future users of IDS models are advised to conduct simulations tailored to their study to gauge how well the model likely performs in their case. In addition, any extra information about availability should be incorporated in the model, such as data from multiple observers (as in mark-recapture distance sampling; Borchers et al. 1998), replicated surveys (Chandler et al. 2011), time-of-detection and time-removal data (Farnsworth et al. 2005; Alldredge et al. 2007; Solymos et al. 2013, Amundson et al. 2014). Alternatively, availability parameters may be estimated from altogether different data types, such as recordings of individual singing behavior, or perhaps even taken from the literature. We note that Solymos et al. (2013) had good success with the integration of time-removal and distance sampling data, but in a simpler model that did not involve estimation of a detection function for the time-removal data.

Most distance sampling models, including our IDS models, assume that survey sites are placed randomly in the study area. However, in our case study, many surveys were done along roadsides, many of which were logging roads within woodlands (Appendix S2: Section S3). We assume that American Robin distribution was unaffected by the vicinity of these roads and our observations of them being distributed well away from roads in the Finley Refuge where we sampled an off-road grid supports that assumption. Furthermore, canopy cover, one of our important environmental covariates, helps to account for the presence of roads and road size as larger or denser roads at a survey location decrease canopy cover. The use of appropriate

environmental covariates, or the intentional inclusion of off-road surveys, should be considered so that effects of roadside v. off-road counts can be evaluated. We also caution that effects of proximity to roads may not affect the distribution of all species equally.

We can envision at least four major extensions to the IDS models described in this paper. First is the accommodation of survey sites included in the dataset which were sampled using multiple protocols. This induces a dependence that must be addressed in the construction of the joint likelihood. Second, IDS models could be developed for other survey geometries, such as line transects or search-encounter designs (Royle et al. 2014, Mizel et al. 2018). Third, allowing for open populations and demographic processes (Kéry & Royle, 2021: chapters 1 and 2) will be an important extension that may open up avenues for truly large-scale demographic models; see also Appendix S1. Fourth, additional data types may be incorporated in the model, such as opportunistic data conceptualized as point patterns (Farr et al. 2021), time-to-detection data (Strebel et al. 2021), aggregated counts (Schmidt et al. 2022), and data from autonomous recording units (ARUs; Doser et al. 2021). For instance, IDS models may be beneficial for ARU data by allowing estimation of the "listening range" of these devices under widely varying conditions, while additionally exploiting the information on singing rate contributed by the ARU data.

In summary, we believe that IDS models can improve analyses of widely available simple PC and DND data obtained in citizen-science schemes, as well as the increasing amount of ARU data in contemporary biodiversity surveys. IDS models may serve as a keystone of the formal, model-based unification of various data types both from designed and less-designed to even design-free surveys, to great mutual benefit. We find it fascinating to see how DS and simple PC or DND data both contribute two essential pieces of information towards the full IDS model: DS data contain most information about the detection function, while the heterogeneity in survey duration commonly found in simple PC/DND data enables estimation of the availability process. This neatly illustrates the fact that the future of biodiversity monitoring arguably lies in a combination of both designed surveys and carefully chosen citizen-science schemes.

**Acknowledgements:** The Bob and Phyllis Mace Professorship to WDR supported the Oregon 2020 project, while Reto Burri commented on an earlier draft. We thank two reviewers of an earlier draft of this paper, as well as the associate editor, for extremely helpful comments. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

**Conflicts of interest:** We declare no conflicts of interest.

## References

Alldredge, M.W., Simons, T.R., Pollock, K.H. 2007. Factors affecting aural detections of songbirds. *Ecological Applications*, 17, 948–955.

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. *Auk*, 131, 476–494. a

Balmer, D.E., Gillings, S., Caffrey, B.J., Swann, B., Downie, I.S., Fuller, R.J. 2013. *Bird Atlas 2007–11: The Breeding and Wintering Birds of Britain and Ireland*. Thetford: BTO Books.

Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A. 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, 58, 540–547.

Borchers, D.L., Efford, M.G. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, 64, 377–385.

Borchers, D.L., Zucchini, W., Fewster, R.M. 1998. Mark-recapture models for line transect surveys. *Biometrics*, 54, 1207–1220.

Buckland, S.T., Rexstad, E.A., Marques, T.A., Oedekoven, C.S. 2015. *Distance sampling: methods and applications.* Springer, Cham, Switzerland.

544 Chandler, R.B., Royle, J.A., King, D.I., 2011. Inference about density and temporary emigration in
545 unmarked populations. *Ecology*, 92, 1429–1435.
546 Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R., Hooten, M.B. 2018. A guide to Bayesian
547 model checking for ecologists. *Ecological Monographs*, 88, 526–542.
548 Darras, K.F.A., Yusti, E., Huang, J.C.-C, Zemp, D.-C., Kartono, A.P., Wanger, T.C. 2021. Bat
549 point counts: A novel sampling method shines light on flying bat communities. *Ecology and*
550 *Evolution*, 11, 17179-17190
551 Diefenbach, D.R., Marshall, M.R., Mattice, J.A., Brauning, D.W. 2007. Incorporating availability
552 for detection in estimates of bird abundance. The Auk, 124, 96–106.
553 Dorazio, R.M., Jelks, H.L., Jordan, F. 2005. Improving removal-based estimates of abundance by
554 sampling a population of spatially distinct subpopulations. *Biometrics*, 61, 1093–1101.
555 Dorazio, R.M. 2012. Predicting the geographic distribution of a species from presence-only data
556 subject to detection errors. *Biometrics*, 68, 1303–1312.
557 Dorazio, R.M. 2014. Accounting for imperfect detection and survey bias in statistical analysis of
558 presence-only data. *Global Ecology and Biogeography*, 23, 1472–1484.
559 Doser, J.W., Finley, A.O., Weed, A.S., Zipkin. E.F. 2021. Integrating automated acoustic
560 vocalization data and point count surveys for estimation of bird abundance. *Methods in*
561 *Ecology and Evolution*, 12, 1040–1049.
562 Elith, J., Kearney, M. Phillips, S. 2010. The art of modelling range-shifting species. *Methods in*
563 *Ecology and Evolution*, 1, 330-342.
564 Farnsworth, G.L., Nichols, J.D., Sauer, J.R., Fancy, S.G., Pollock, K.H., Shriner, S.A., Simons,
565 T.R. 2005. *Statistical Approaches to the Analysis of Point Count Data: A Little Extra*
566 *Information Can Go a Long Way*. USDA Forest Service Gen. Tech. Rep. PSW-GTR-191.
567 Farr, M.T., D.S. Green, K.E. Holecamp, E.F. Zipkin. 2021. Integrating distance sampling and
568 presence-only to estimate species abundance. *Ecology*, 102, e03204.

Fiske, I., Chandler, R. 2011. unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, 43, 1–23.

Kellner, K.F., A.D. Smith, J.A. Royle, M. Kéry, J.L. Belant, R.B. Chandler. 2023. The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology. *Methods in Ecology and Evolution*, 14, 1408–1415.

Hallman, T.A., Robinson, W.D. 2020. Comparing multi- and single-scale species distribution and abundance models built with the boosted regression tree algorithm. *Landscape Ecology*, 35, 1161–1174.

Hostetter, N.J., Gardner, B., Sillett, T.S., Pollock, K.H., Simons, T.R. 2019. An integrated model decomposing the components of detection probability and abundance in unmarked populations. *Ecosphere*, 10(3).

Kellner, K. 2016. jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses. R package version 1.4.4. CRAN.R-project.org/package=jagsUI

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. *AHM1—Modeling distribution, abundance and species richness using R and BUGS. Volume 1: Prelude and Static Models*. Elsevier / Academic Press.

Kéry, M., Royle, J.A. 2021. *AHM2—Modeling distribution, abundance and species richness using R and BUGS. Volume 2: Dynamic and Advanced Models*. Elsevier / Academic Press.

Kéry, M., Royle, J.A., Hallman, T., Robinson, D.W., Strebel, N., Kellner, K.F. 2024. Simulation and analysis code for "Integrated distance sampling models for simple point counts" (v1.0). Zenodo. https://doi.org/XXXXXXXX.

Lele, S.R., Moreno, M., Bayne, E. 2012. Dealing with detection error in site occupancy surveys: what can we do with a single survey? *Journal of Plant Ecology*, 5, 22–31.

LEMMA. 2014. Landscape Ecology, Modeling, Mapping, and Analysis LEMMA. GNN Structure Maps. https://lemma.forestry.oregonstate.edu/data/structure-maps. Accessed 6 Sep 2016

595 Marques, T.A., Thomas, L., Fancy, S.G., Buckland, S.T., 2007. Improving estimates of bird density
596 using multiple-covariate distance sampling. The Auk, 124(4), 1229-1243.
597 Marsh, H., Sinclair, D.F. 1989. Correcting for visibility bias in strip transect aerial surveys of
598 aquatic fauna. *Journal of Wildlife Management*, 53, 1017–1024.
599 Miller, D.A.W., Pacifici, K., Sanderlin, J.S., Reich, B.J. 2019. The recent past and promising future
600 for data integration methods to estimate species' distributions. *Methods in Ecology and*
601 *Evolution*, 10, 22–37.
602 Mizel, J.D., Schmidt, J.H., Lindberg, M.S. 2018. Accommodating temporary emigration in spatial
603 distance sampling models. *Journal of Applied Ecology*, 55, 1456–1464.
604 Nichols, J.D., Hines, J.E., Sauer, J.R., Fallon, F.W., Fallon, J.E., Heglund, P.J. 2000. A double-
605 observer approach for estimating detection probability and abundance from point counts.
606 *Auk*, 117, 393–408.
607 Nichols, J.D., Thomas, L., Conn, P.B. 2009. Inferences about landbird abundance from count data:
608 recent advances and future directions. pp. 201–235 in D.L. Thomson, E.G. Cooch, M.J.
609 Conroy (eds.) *Modeling demographic processes in marked populations*. Springer.
610 Oedekoven, C.S., Laake, J.L., Skaug, H.L. 2015. Distance sampling with a random scale detection
611 function. *Environmental and Ecological Statistics*, 22, 725–737.
612 Oregon Spatial Data Library. 2017. Oregon 10m Digital Elevation Model (DEM).
613 http://spatialdata.oregonexplorer.info/geoportal/details;id=7a82c1be50504f56a9d49d13c7b4
614 9aa. Accessed 30 Nov 2015
615 Pacifici, K., B.J. Reich, D.A.W. Miller, B.S. Pease. 2019. Resolving misaligned spatial data with
616 integrated species distribution models. *Ecology*, 100, e02709.
617 Péron, G., Garel, M. 2019. Analzying patterns of population dynamics using repeated population
618 surveys with three types of detection data. *Ecological Indicators*, 106, 105546.

Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop in Distributed Statistical Computing (DSC 2003) (eds K. Hornik, F. Leisch& A. Zeileis), 1–10, TU Vienna, Austria.

R Core Team. 2019. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Robinson, W.D., T.A. Hallman, J.R. Curtis. 2020. Benchmarking the avian diversity of Oregon in an era of rapid change. *Northwestern Naturalist*, 101, 180–193.

Robinson, W.D., T.A. Hallman, R.A. Hutchinson. 2021. Benchmark bird surveys help quantify counting accuracy in a citizen-science database. *Frontiers in Ecology and Evolution*, 9, https://doi.org/10.3389/fevo.2021.568278.

Rosenstock, S.S., Anderson, D.R., Giesen, K.M., Leukering, T., Carter, M.F. 2002. Landbird counting techniques: current practices and an alternative. *Auk*, 119, 46–53.

Royle, J.A. 2004. N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60, 108–115.

Royle, J.A., Chandler, R.B., Sollmann, R., Gardner, B., 2014. *Spatial Capture-Recapture*. Academic Press.

Royle, J.A., Dawson, D.K., Bates, S. 2004. Modeling abundance effects in distance sampling. *Ecology*, 85, 1591–1597.

Royle, J.A., Nichols, J.D. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology*, 84, 777–790.

Sauer, J.R., Pardieck, K.L., Ziolkowski Jr, D.J., Smith, A.C., Hudson, M.A.R., Rodriguez, V., Berlanga, H., Niven, D.K., Link, W.A. 2017. The first 50 years of the North American breeding bird survey. *The Condor: Ornithological Applications*, 119, 576–593.

Schaub, M., Kéry, M. 2022. *Integrated Population Models — Theory and ecological applications with R and JAGS.* Elsevier, Academic Press.

644 Schmidt, J.H., Wilson, T.L., Thompson, W.L., Mangipane, B.A. 2022. Integrating distance
645 sampling survey data with population indices to separate trends in abundance and temporary
646 immigration. *J. Wildlife Management*, 86, e22185.
647 Sólymos, P., Matsuoka, S.M., Bayne, E.M., Lele, S.R., Fontaine, P., Cumming, S.G., Stralberg, D.,
648 Schmiegelow, F.K.A., Song, S.J. 2013. Calibrating indices of avian density from non-
649 standardized survey data: making the most of a messy situation. *Methods in Ecology and*
650 *Evolution*, 4, 1047–1058.
651 Strebel, N., Fiss, C.J., Kellner, K.F., Larkin, J.L., Kéry, M., Cohen, J. 2021. Estimating abundance
652 based on time-to-detection data. *Methods in Ecology and Evolution*, 12, 909–920.
653 Stoudt, S., de Valpine, P., Fithian, W. 2023. Nonparametric identifiability in species distribution
654 and abundance models: Why it matters and how to diagnose a lack of it using simulation.
655 *Journal of Statistical Theory and Practice* 17:39. doi.org/10.1007/s42519-023-00336-5
656 Sullivan, B.L., C.L. Wood, M.J. Iliff, R.E. Bonney, D. Fink, S. Kelling. 2009. eBird: A citizen-
657 based bird observation network in the biological sciences. *Biological Conservation*, 142,
658 2282–2292.
659 United States Geological Survey. 2011. National Gap Analysis Project.
660 https://gapanalysis.usgs.gov/gaplandcover/data/download/. Accessed 7 Dec 2015
661 Wyatt, R.J. 2002. Estimating riverine fish population size from single- and multiple-pass removal
662 sampling using a hierarchical model. *Can. J. of Fish. and Aquatic Sciences*, 59, 695–706.
663 Zipkin, E.F., Rossman, S., Yackulic, C.B., Wiens, J.D., Thorson, J.T., Davis, R.J., Grant, E.H.C.
664 2017. Integrating count and detection–nondetection data to model population dynamics.
665 *Ecology*, 98(6), 1640-1650.

**Figure captions**

Fig. 1: Simulation-based validation of two integrated distance sampling (IDS) models (Simulation 1). Left: Model IDS1 (=DS + PC data), right: Model IDS2 (=DS + DND data); see main text for details. Top: estimation error in detection function sigma ($\sigma$) in the DS data (n = 250 sites); middle: estimated (with 95% CRIs) *vs.* true value of $\sigma$ in the PC and the DND data sets (n = 1000 sites); bottom: estimation error in the latent site-level abundances (*N*) in the PC and the DND data (mean/sd of simulated true abundance: 79/9). Red denotes truth or absence of estimation error, dashed blue shows mean of estimates. Sample size in both simulations is 1000 data sets. See also Appendix S2: Figure S1 and Tables S1-S2.

Fig. 2: Another simulation-based validation of IDS1 combining DS and PC data (Simulation 2). Left: Sampling distributions of intercept and slope estimates for detection function parameters with independent effects in the distance sampling (top) and the point count (bottom) parts of the data (Simulation 2a). Right: Intercept and slope estimates for detection function parameters with independent effects in the distance sampling (top) and the point count (bottom) parts of the data, when the same covariate has also an effect on density (Simulation 2b). Red denotes truth, dashed blue shows mean of estimates. Sample size in both simulations is 1000 data sets. See also Appendix S2: Table S3.

Fig. 3: Sampling distributions of estimators of density (intercept and slope of a continuous covariate, shared between distance sampling (DS) and simple point count (PC) data), and of detection function sigma ($\sigma$) for the DS and the PC parts of the data (Simulation 2). Throughout, sample size for the simple PCs is 200 and true values are indicated with dashed red lines. Each individual boxplot summarizes between 515 and 998 data sets that resulted in valid estimates, see also Appendix S2: Table S4. Note that more variable boxplots are indicative of higher RMSEs.

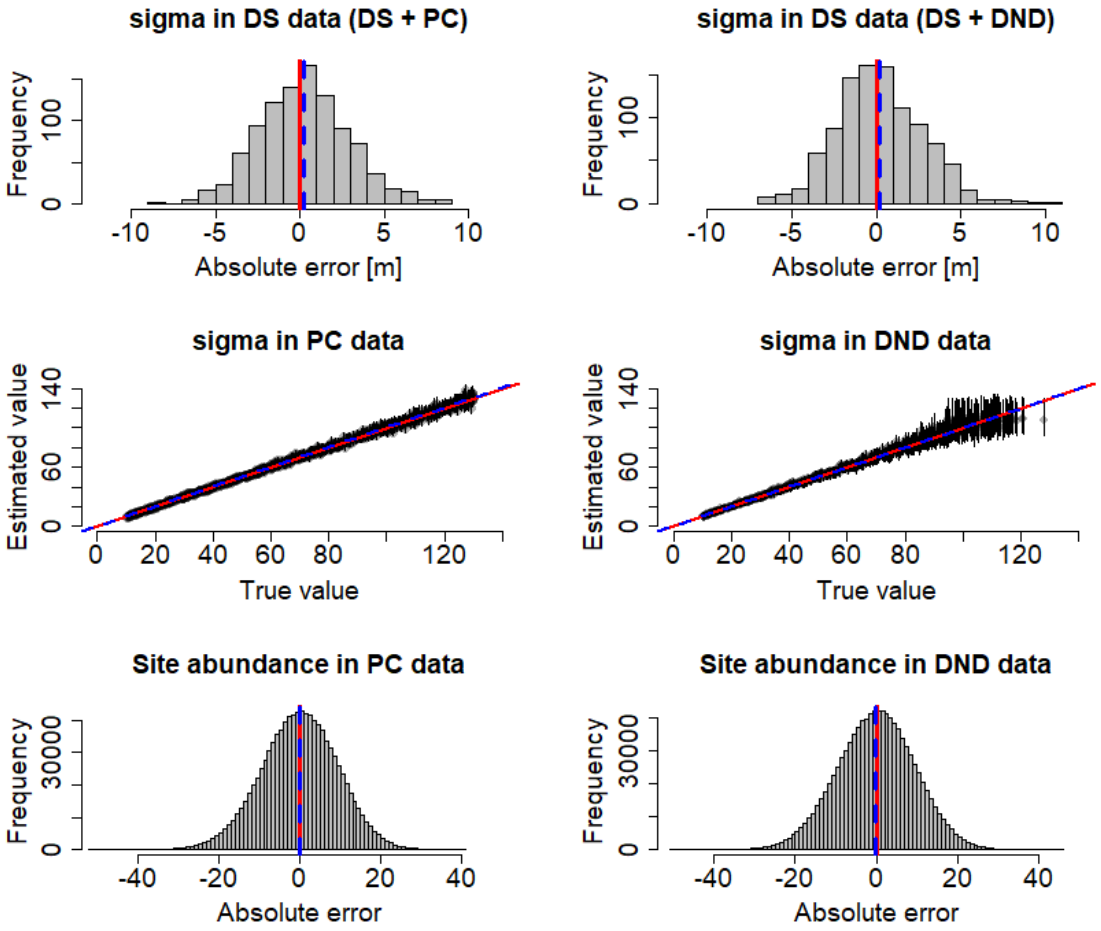
Fig. 4: Sampling distributions of estimators of density (lambda, $\lambda$) and of activity/singing rate (phi, $\phi$) in an IDS model with availability fit to data from 3000 DS sites, plus 1000, 3000, or 6000

691 PC sites added (Simulation 3, n = 930, 866, and 997 analyses that did not produce numerical
692 failures). Red denotes truth or absence of estimation error, dashed blue shows mean of valid
693 estimates. See also Appendix S2: Table S5.

694 Fig. 5: Estimated density of American Robin (individuals per 1 $km^2$) in Benton and Polk counties,
695 Oregon, based on breeding season observation data from 2011 to 2017.

696 Figure 1

697




698

699

700

701

Figure 2

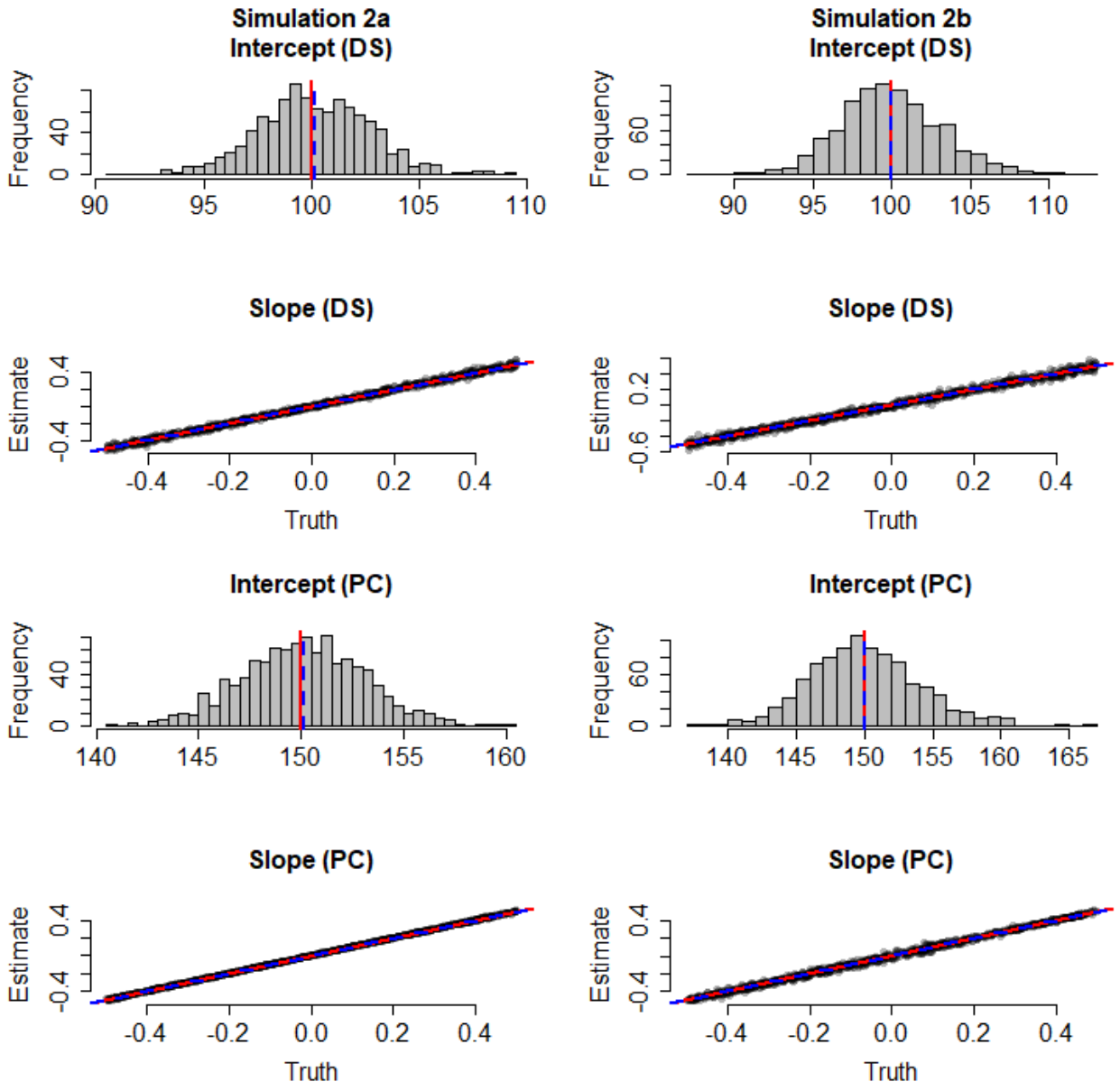
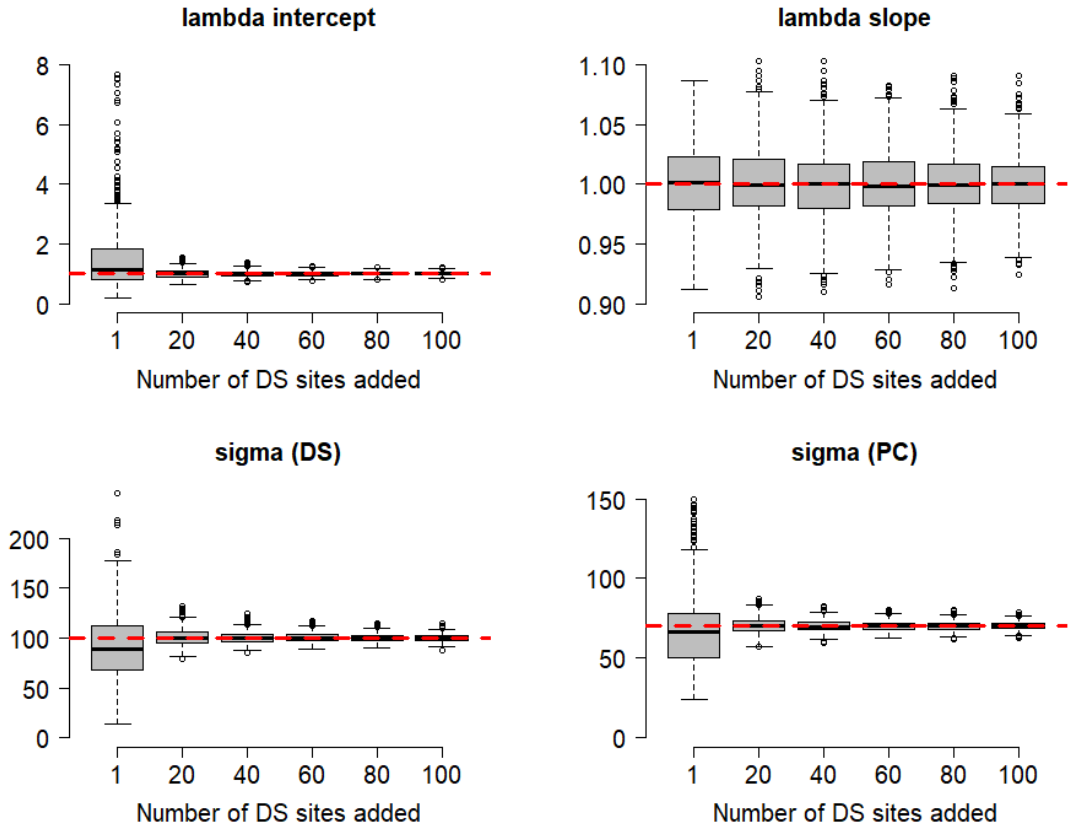
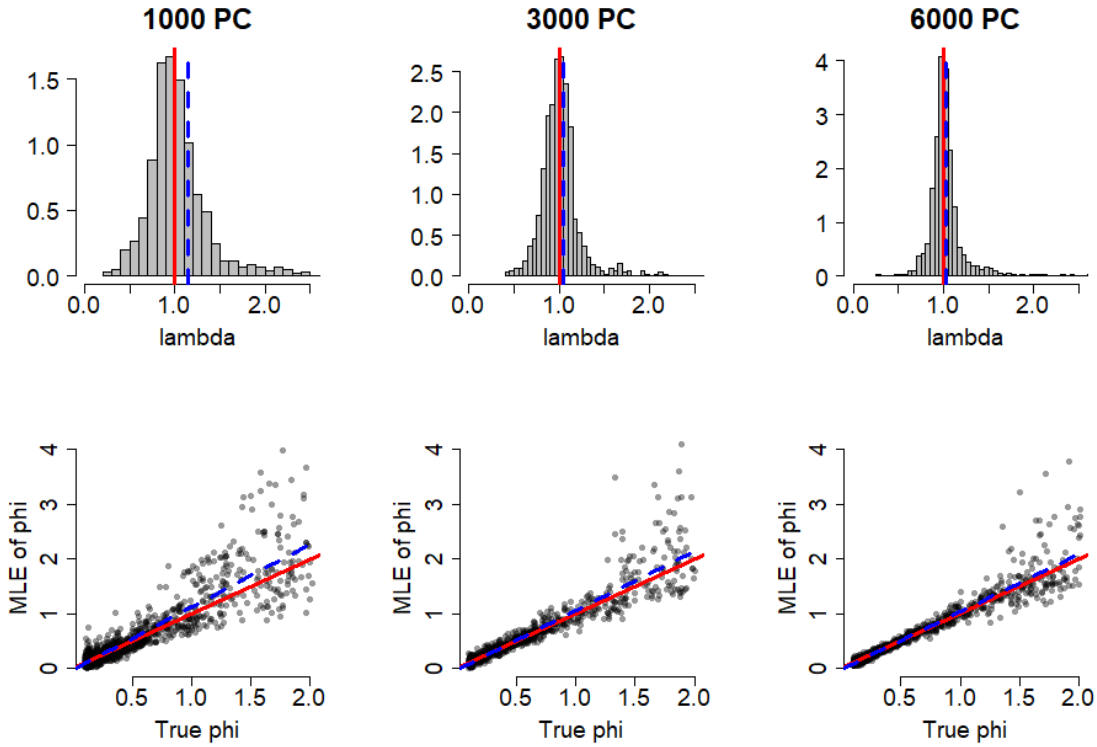

Simulation 2a
Intercept (DS)
Frequency
0
40
90
95
100
105
110
Simulation 2b
Intercept (DS)
Frequency
0
60
90
95
100
105
110
Slope (DS)
Estimate
-0.4
0.4
-0.4
-0.2
0.0
0.2
0.4
Truth
Slope (DS)
Estimate
-0.6
0.2
-0.4
-0.2
0.0
0.2
0.4
Truth
Intercept (PC)
Frequency
0
40
140
145
150
155
160
Intercept (PC)
Frequency
0
60
140
145
150
155
160
165
Slope (PC)
Estimate
-0.4
0.4
-0.4
-0.2
0.0
0.2
0.4
Truth
Slope (PC)
Estimate
-0.4
0.4
-0.4
-0.2
0.0
0.2
0.4
Truth

Figure 3


lambda intercept
8
6
4
2
0
1
20
40
60
80
100
Number of DS sites added
lambda slope
1.10
1.05
1.00
0.95
0.90
1
20
40
60
80
100
Number of DS sites added
sigma (DS)
200
150
100
50
0
1
20
40
60
80
100
Number of DS sites added
sigma (PC)
150
100
50
0
1
20
40
60
80
100
Number of DS sites added

711 Figure 4

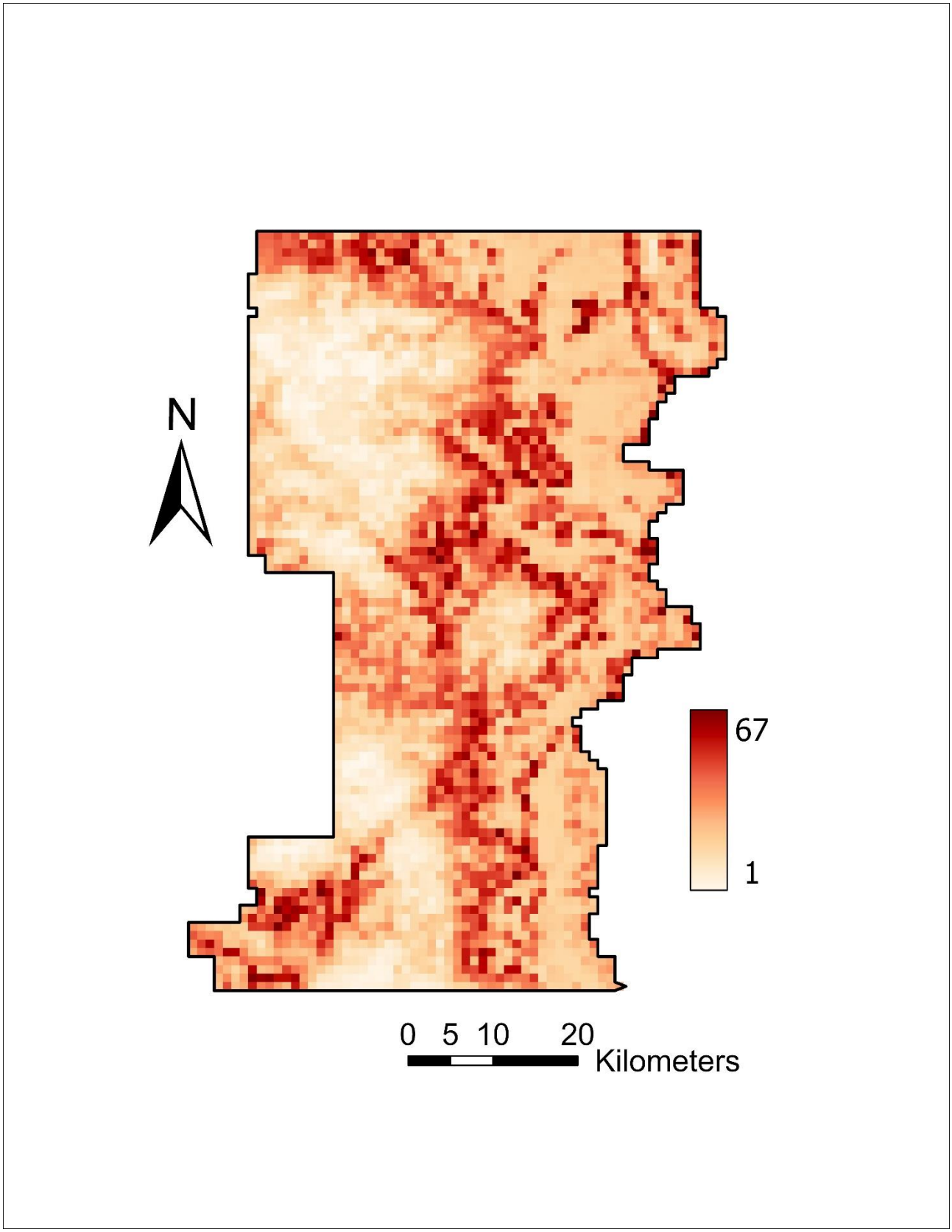



712

713

714

715 Figure 5




716

717