




Research Article

Scaling Occupancy Estimates up to Abundance for Wolves

GLENN E. STAUFFER,¹ *Office of Applied Sciences, Wisconsin Department of Natural Resources, 107 Sutliff Ave, Rhinelander, WI 54501, USA*
 NATHAN M. ROBERTS,² *Office of Applied Sciences, Wisconsin Department of Natural Resources, 107 Sutliff Ave, Rhinelander, WI 54501, USA*
 DAVID M. MACFARLAND, *Office of Applied Sciences, Wisconsin Department of Natural Resources, 107 Sutliff Ave, Rhinelander, WI 54501, USA*
 TIMOTHY R. VAN DEELEN , *Department of Forest and Wildlife Ecology, University of Wisconsin, 217 Russell Labs, 1630 Linden Dr, Madison, WI 53706, USA*

ABSTRACT Management of wildlife populations often requires reliable estimates of population size or distribution. Estimating abundance can be logistically difficult, and occupancy models have been used as a less expensive proxy for abundance estimation. Another alternative is to use independent estimates of home-range size and mean group size to directly scale occupancy estimates up to abundance. We used simulations to explore when scaling occupancy up to abundance is reliable, and as an example we applied an occupancy approach to estimate abundance of wolves (*Canis lupus*) from roadside snow-tracking surveys in northern Wisconsin, USA, in 2016 and 2018. Estimates of wolf abundance were plausible and compared favorably with independent estimates produced by territory mapping, and snow-tracking data requirements were lower than for territory mapping. Simulation results suggested that reasonable abundance estimates could be obtained under some conditions but also that severe positive bias could result under other conditions, especially when populations were small and dispersed, home range size was small, and areal sampling units were large. Positive bias in abundance estimates occurs because of closure assumption violations when tracks from a single wolf or pack are detected in >1 sample unit, and the sum of the sample unit areas where tracks were detected exceed the sum of the home range areas. Bias was minimized when sampling units were small relative to home range size or when sampling units were route segments that approximate point sample units, and when home ranges were highly aggregated. We conclude that, although caution is warranted when scaling occupancy estimates up to abundance, scaled occupancy models can provide feasible and reliable estimates of abundance, assuming home range size and mean group size are accurately known or estimated, sampling units are appropriately chosen, and covariates that aggregate home ranges can be used to accurately predict occupancy probability. © 2021 The Authors. *The Journal of Wildlife Management* published by Wiley Periodicals LLC on behalf of The Wildlife Society.

KEY WORDS estimator bias, monitoring, multi-scale occupancy, simulation, territory mapping, Wisconsin, wolves.

Estimating population size or trend is often a necessary and primary objective of wildlife population monitoring programs. In some cases where collection of replicated count data or capture-recapture data is infeasible or prohibitively expensive, management objectives can be met by instead estimating the proportion of an area occupied by a species. Estimating occupancy requires only spatially or temporally replicated detection–non-detection data, which often are less expensive to collect than counts or capture-recapture

data. Occupancy models (MacKenzie et al. 2002, Tyre et al. 2003) are now widely used to estimate distribution or as a proxy for abundance for a wide variety of taxa (Durso et al. 2011, Falke et al. 2012, Adams et al. 2013, Chen et al. 2013, Rich et al. 2016), or scaled up to directly estimate abundance (Rich et al. 2013, Ausband et al. 2014). Occupancy models account for the common situation where a species may be present but undetected at a specific survey site during a specific survey occasion, and thus help prevent spurious conclusions about habitat relationships that result from confounding species absences and species non-detection (Tyre et al. 2003, Gu and Swihart 2004).

Occupancy (ψ) is a useful metric, but interpretation of ψ depends on the specifics of a study design (Efford and Dawson 2012, Latif et al. 2016, Steenweg et al. 2018). With discrete sample units, ψ can be interpreted as the instantaneous probability of patch occupancy (MacKenzie et al. 2002), but often, especially for temporally replicated surveys in continuous habitat, ψ represents cumulative or

Received: 17 August 2020; Accepted: 19 May 2021

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

¹E-mail: Glenn.Stauffer1@wisconsin.gov

²Current address: Division of Mathematics, Science, and Engineering, College of the Ozarks, 100 Opportunity Ave, Point Lookout, MO 65726, USA

asymptotic probability of patch occupancy (Tyre et al. 2003, Efford and Dawson 2012). Failure to carefully define occupancy can lead to misinterpretation and inferential errors (Efford and Dawson 2012, Hayes and Monfils 2015). Moreover, when individuals are distributed randomly on the landscape, expected cumulative occupancy probability, $E(\psi)$, is a function of the ratio of home range size to sample unit size (R) and population density, and for a given population density, increasing R increases $E(\psi)$, and numerous combinations of home range size, sample unit size, and density can result in identical estimates of cumulative occupancy (Efford and Dawson 2012: figure 2).

Given closure assumption violations and consequent different interpretations for occupancy probability, special care is warranted when scaling occupancy up to abundance or when using occupancy as an index for abundance or density. Indices rarely scale linearly with abundance beyond some narrow range (MacFarland and Van Deelen 2011, Steenweg et al. 2018), and ψ has an obvious asymptote at 1 regardless of any possible further increases in population density. MacKenzie et al. (2006) suggested that when sample unit size is chosen to match territory size of territorial species, then the number of occupied units is virtually equivalent to the number of territorial individuals or pairs. Although not stated explicitly, in continuous habitat where sample units are not closed to temporary emigration, that correspondence likely depends on the temporal scale of sampling (Efford and Dawson 2012, Steenweg et al. 2018).

If the number of occupied sites adequately describes the distribution of home ranges (i.e., ψ is appropriately defined and interpreted, and is unbiased), then population abundance can be derived as $N = (\psi A \bar{x}) / \bar{h}$ where ψ is the probability of occupancy in a sample unit, A is the total area of the sample units (or the area study, if sample units cover the entire study area), \bar{h} is the mean home range size during the sampling period, and \bar{x} is the mean group, or pack size. This approach is currently used to estimate abundance of wolves (*Canis lupus*) in Montana and Idaho, USA (Rich et al. 2013, Ausband et al. 2014). Scaling occupancy up to abundance in this way obviously entails important assumptions about territory size, group size, and group cohesion (Rich et al. 2013, Ausband et al. 2014, Latham et al. 2014).

Assuming that territory size and group size are known or can be properly estimated, scaling occupancy to abundance requires consideration of 2 questions: does ψ represent instantaneous or cumulative occupancy and how can potential bias from expected closure assumption violations be minimized or mitigated? Scaling occupancy as $N = (\psi A \bar{x}) / \bar{h}$ implicitly defines ψ as cumulative occupancy, or the proportion of the landscape included in a home range. Because closure cannot be assured, and features that mitigate or minimize bias are not entirely controllable, it is also important to identify scenarios where consequences of violations are minimized, or where an occupancy abundance approach might be inappropriate. For example, selection of sample units is controllable, and large sample units should exacerbate consequences of closure violations. Conversely,

landscape features and territory distribution are not controllable, but if landscape features influence distribution such that home ranges are clustered, then closure violations should be less consequential than if home ranges are widely scattered.

It is worth considering whether there are situations where a closure assumption could be relaxed. One possible approach is to collect data in a robust design framework and fit them to a dynamic occupancy model (MacKenzie et al. 2003). This approach allows occupancy status in sample units to change between primary periods and considers initial occupancy probability (ψ_0) to be an unbiased estimate of instantaneous occupancy that captures the distribution of home range centers. Spatial subunits could serve as replicates, assuming the target species is mobile enough that all subunits could conceivably be occupied during a single visit (Kendall and White 2009). This assumption is problematic in a continuous landscape, where mobile animals could be detected in adjacent sample units within a single survey occasion. In this case, ψ_0 probably is still closer to cumulative occupancy than instantaneous occupancy, especially when detection relies on persistent cues such as tracks or scat. The equal availability assumption also is violated when a home range intersects with only some subunits of >1 sample unit.

Another possible approach is to use multiscale models, where data are collected at temporal or spatial subsamples nested within site or survey occasion (Mordecai et al. 2011, Pavlacky et al. 2012, Whittington et al. 2015). In this approach, ψ represents overall occupancy, and θ represents local occupancy at a spatial subunit, or availability at a site during a specific survey occasion. Efford and Dawson (2012) suggest that, even when replication at a site is spatial rather than temporal, a temporal context should be considered, where ψ represents asymptotic occupancy and $\psi\theta$ represents instantaneous occupancy. The aforementioned caveats about highly mobile individuals also apply to multi-scale models.

In Wisconsin, USA, wolf abundance has historically been estimated by territory mapping and enumeration of presumed wolf packs, and pack territories are inferred from roadside snow-tracking surveys, evidence of breeding, telemetry locations, and aerial observations (Wydeven et al. 2009). Assignment of observed tracks into pack territories is somewhat circular because assignments are made based on presumed territorial locations of the different packs, but territories are inferred in part from the location of the observed tracks. This enumeration is regarded as an annual minimum over-winter count, with bounds defined by cumulative uncertainty about individual pack sizes.

Although territory mapping has produced plausible estimates of wolf abundance in Wisconsin, estimates rely on somewhat subjective pack assignments, and likely deviate from true abundance to an unknown and possibly variable degree. Also, although estimates are presented as a narrow range, uncertainty is not rigorously evaluated. Precise estimates are desirable, but management decisions should

reasonably account for all sources of uncertainty, rather than treating estimates with unknown bias and uncertainty as truth (Yoccoz et al. 2001). Accordingly, there is motivation to develop an estimator that accounts for incomplete detection, provides a rigorous estimate of uncertainty, and is less reliant on uncertain pack assignments.

We conducted a simulation study to investigate bias in occupancy and abundance approaches for estimating abundance of wolves for various realistic scenarios in a continuous landscape where closure assumption violations routinely occur. Our objectives were to use simulations to identify scenarios where scaling occupancy estimates to abundance resulted in minimal bias, and to illustrate a scaled occupancy approach for estimating wolf abundance from snow tracking data in northern Wisconsin, during winter 2016–2017 and 2018–2019. We evaluated a variety of simulation scenarios, with varying densities, distribution patterns, and sampling and modeling strategies, and we predicted minimal bias in scenarios where closure assumption violations were minimal or relatively inconsequential (e.g., when home ranges are aggregated and the landscape or portions thereof are essentially saturated) and when sample unit sizes are appropriately chosen.

STUDY AREA

We analyzed snow-tracking data from 2016–2019 collected from Wolf Management Units 1 and 3 in northern-western Wisconsin, and from the extent of the known wolf range in northern and central Wisconsin (Fig. 1). The known wolf range comprised approximately 91,000 km², although the precise extent of the surveyed area varied annually, with elevation varying from 177 m at Lake Michigan to 595 m at Timms Hill. The area was approximately 68% forested, 19.3% cultivated or grassland, 4.3% residential or otherwise developed, and 3% emergent wetlands. Northern Wisconsin was glaciated and dominated by mixed forest and forested wetlands, with common trees including pines (*Pinus* spp.), maples (*Acer* spp.), birch (*Betula* spp.), and oaks (*Quercus* spp.), whereas central Wisconsin was largely unglaciated and agricultural, except for a sparsely populated area of about approximately 7,100 km², known as the Central Forest Region, largely disjunct from the forests of northern Wisconsin (Theil et al. 2009). Primary land uses throughout the region were forestry, agriculture, and tourism. Common large mammalian species included white-tailed deer (*Odocoileus virginianus*), American black bear (*Ursus americanus*), coyote (*Canis latrans*), wolf, red fox (*Vulpes vulpes*), and beaver (*Castor canadensis*). Summers (Jun–Aug) were generally warm and humid, with average summer temperatures of approximately 18°C, and a freeze-free season of approximately 100 days in the north. Winters (Dec–Mar) were cold and snowy, with average winter temperatures of approximately –11°C, and average annual snow cover persistence from 140 days in the north to approximately 85 days in the south. Spring (Apr–May) and fall (Sep–Nov) were typically cool, with variable precipitation, and with occasional late or early frosts.

METHODS

Simulation 1: Areal Sample Units

We simulated data for a 180 × 180-km landscape, with an additional buffer 2 times the diameter of the simulated mean home range size to allow wolf movement in and out of the study area. We placed wolves ($N = 100$ or 400) on the landscape in packs with mean size 3.8 and maximum pack size 8. About 15% of packs had only a single wolf (i.e., loners). Wolves in the same pack had identical home range centers, and the simulation drew home range sizes from a normal distribution with a mean of 100 km² or 324 km² with a standard deviation of 10. Thus, in our simulations with large population and large home range sizes, the landscape was nearly saturated with wolf territories. We selected home range centers by gridding the landscape with a resolution equal to μ/c , where μ was mean home range size and $c = 1.15$ is a multiplier allowing some home range overlap. Realized c is dependent on home range size and the number and spatial arrangement of home ranges on the landscape and may exceed 1.15 because simulated home ranges are circular rather than square. We simulated 100 replicates of each abundance-home range combination to determine the value of c to use for inflating abundance estimates. We assigned selection probabilities to the gridded study area, with some spatial heterogeneity deliberately introduced, and selected home range centers without replacement.

Once home ranges were established, we imposed a new sampling grid, with either 100 km² or 324 km² resolution, with 4 subunits per cell. We specified availability for detection for different scenarios based on 1 of 6 wolf movement models, which differed in the extent to which packs

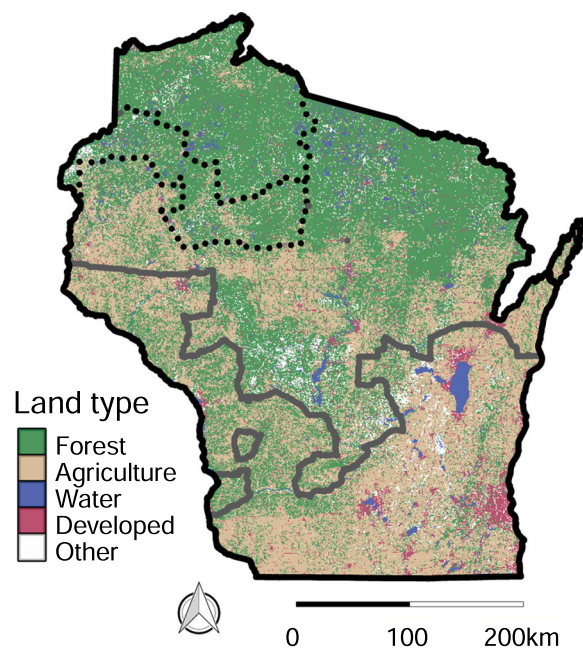


Figure 1. Wolf Management Units 1 and 3 (black dotted lines) and total extent of surveyed wolf range (gray solid lines) in Wisconsin, USA, in 2016–2019. Habitat classifications are derived from 2016 National Land Cover Data (Yang et al. 2018).

were cohesive and spatio-temporally excluded other packs. In the most realistic model (model 1), used in most of our scenarios, we intersected the sampling grid with home ranges and assigned availability probabilities for pack k (or individuals in pack k) in the resulting home range-sample subunit fragments as $\text{logit}(\theta_j^k) = \alpha + \log(ppn_j^k)$, where α defines baseline availability when the entire home range k is contained inside a single sample unit, and where ppn_j^k is the proportion of the home range contained in the j th fragment of pack k , and $\alpha = 0.62$ or 2.2 , corresponding to $\theta_j^k = 0.65$ or 0.9 , respectively, when $ppn_j^k = 1$. Availability of a pack (when movements of individuals in a pack were completely dependent: model 1a) or individual (when individuals could move and be detected independent of other pack members: model 1b) within home range fragment k in fragment j was then defined as $u_j^k \sim \text{Bin}(\theta_j^k)$. Model 1b simulates a scenario with little to no pack cohesion other than sharing the same home range. Thus packs (or individuals) were available with some probability in each subunit within their home range (which might be part of >1 primary sample unit) in each survey period. This is reasonable and realistic if availability is interpreted to be the production of sign such as scat or tracks that may persist long enough to be detected in >1 location during the same survey occasion. In this case, the resulting estimate should be close to cumulative, not instantaneous occupancy probability. We also simulated availability with three alternative models. The first is as above, except that availability of packs (model 2a) or individuals (model 2b) was modeled with a multinomial model $\{u_1^k, u_2^k, \dots, u_J^k\} \sim \text{MN}(1, \{\theta_1^k, \theta_2^k, \dots, \theta_J^k\})$, where J is the number of home range fragments available in the territory of individual (or pack) k . Thus, packs or individuals were available in only one home range fragment in any given survey occasion. This model violates the equal availability assumption of spatial replication but may be realistic if animals move slowly around the home range. In the second alternative (model 3), we first selected an initial pack location randomly from within the home range, subject to the constraint that the location must be ≥ 1 km from other packs. Individuals from the pack were then allowed to deviate from this position based on a $\text{MVN}(0, 25)$ movement kernel. Thus, individuals within a pack could be separated from each other by up to 1 km at each time step, but usually were much closer together. Given the selected sample cell i for each individual, individuals then were assumed to be available (with probability θ_i^k) in all J subunits of cell i , even if some subunits were outside the home range, and were not available in any other sample cell. This is a somewhat unrealistic scenario, but it satisfies the movement and equal availability assumption. Lastly, in the third alternative (model 4), we first selected an initial pack location and subsequent individual locations as described for model 3. But, as with availability model 2, only a single location per occasion was assumed. This model differs from model 2 in that selected locations potentially are constrained by the locations of other packs, and individual members of a pack could, by chance, occupy different sample units or subunits.

In each sample cell, we simulated 0–5 surveys, with most cells having 2–3 surveys. We drew survey effort from a $\text{Beta}(\alpha=5, \beta=2.5)$ distribution and represented effort relative to maximal possible effort (e.g., the length of roads surveyed relative to the maximum length in any sample subunit). Given availability in sample subunit j in grid cell i , we then simulated observations as $y_{ijt} \sim \text{Bernoulli}(p_{ijt})$, where p_{ijt} is detection probability in sample subunit j in grid cell i at time t , $\text{logit}(p_{ijt}) = a_0 + \log(\text{effort}_{ijt})$, and where $a_0 = -0.845$ or 1.099 , respectively, represented relatively low or high detection probability. For the static occupancy model, we aggregated y_{ijt} to $y_{it} = I(\sum_{j=1}^{J_i} y_{ijt})$, where J_i is the number of subunits in sample unit i , and I took the value 1 if the summation was >0 .

When individuals have high movement rates and can be detected in multiple sample units, either over time or within the same sampling occasion, estimated ψ typically represents cumulative use rather than instantaneous occupancy (Efford and Dawson 2012). This should especially be the situation when sample units are adjacent grid cells and an individual's home range could conceivably include portions of >2 grid cells. To try to mitigate this scenario, we conducted a set of simulations where we modeled availability with the most realistic model (model 1) but sampled only in alternating grid cells. We simulated 6 scenarios with abundance = 100 or 400, home range size = 225 km², and sampling grid size = 100, 225, or 324 km². Abundance estimated by $N = (\psi A \bar{x}) / \bar{h}$ then represents only abundance within the sampled units, and must be scaled by A/A_{SA} , where A_{SA} is the size of the sampled area, to estimate the population size. For the simulated cases, where the study area was a uniform square area, this multiplier was exactly 2. We conducted simulations in program R (R Core Team 2018) and fitted models using the R package *runjags* (Denwood 2016).

Occupancy Models

We fitted 3 different occupancy models to the simulated data. The first was a single-season occupancy model:

$$z_i \sim \text{Bernoulli}(\psi)$$

$$y_{it} \sim \text{Bernoulli}(z_i p_{it})$$

$$\text{logit}(p_{it}) = a_0 + \log(\text{effort}_{it}),$$

where z_i is the latent occupancy state of sample cell i . This model should estimate cumulative occupancy probability and potentially result in overestimates of abundance for many scenarios. To scale occupancy up to abundance, we first estimated the area occupied as $\sum_{i=1}^K \psi \times \text{cellsize}_i$, where K was the number of cells, then divided by mean territory size to estimate the number of non-overlapping packs that would fit into the total area occupied (N_{packs}), then multiplied N_{packs} by the mean pack size (μ_{pack}). We accounted for uncertainty by specifying reasonable, empirically informed priors to the

model $\mu_{pack} \sim \text{Gamma}(\alpha = 1,000, \beta = 1,000/\bar{x}_{pack})$, where $\bar{x}_{pack} = 3.8$ was mean pack size estimated from road-side snow-tracking data and aerial pack observations (Wiedenhoeft et al. 2018); and mean home range $(\mu_{HR}) \sim N(\delta_{HR}, SD_{HR})$, where δ_{HR} was a simulation value, and $SD_{HR} = 10$, an empirical value. To account for territory overlap, we inflated N packs with an estimate of home-range overlap (c) as described above.

The second model was a dynamic occupancy model specified by:

$$z_{it} \sim \begin{cases} \text{Bernoulli}(\psi_0), & i = 1 \\ \text{Bernoulli}(z_{it}(1 - \varepsilon) + (1 - z_{it})\gamma), & i > 1 \end{cases}$$

$$y_{ijt} \sim \text{Bernoulli}(z_{it}p_{ijt}),$$

where p_{ijt} was a function of effort, as described above, ε was the extinction probability (probability that an occupied cell becomes unoccupied), and γ was the colonization probability (probability that an unoccupied cell becomes occupied). We scaled up occupancy (ψ_0) to abundance as described above. Because we assumed that the proportion of cells occupied in each time period was constant (even though occupancy of an individual cell can change), we imposed a stationary Markov process (MacKenzie et al. 2006) such that $\varepsilon = \gamma(1 - \psi)/\psi$, which is reasonable, provided that ψ is not too low and γ is not too high.

The third occupancy model was a multiscale model (Mordecai et al. 2011) where

$$z_i \sim \text{Bernoulli}(\psi)$$

$$u_{it} \sim \text{Bernoulli}(z_i\theta)$$

$$y_{ijt} \sim \text{Bernoulli}(u_{it}p_{ijt}),$$

and where p_{ijt} was defined as before, u_{it} was the latent use state in cell i at time t , and θ was the probability that a cell

was being used at time t (not that this definition is different than that described for the data generating models). We scaled up the product $\psi\theta$ (Efford and Dawson 2012) to abundance as described above.

Simulation 2: Linear Sample Units

In the previous section, we generated data generically for areal sample units without describing how such data might be collected. In reality, data often are collected along transects, such as road segments. If defined appropriately, road segments may approximate point sampling units and possibly mitigate biases expected with areal sampling units (Efford and Dawson 2012). We conducted a series of simulations where we used the occupancy model of Crosby and Porter (2018) to generate data from simulated survey routes, and we fitted the simulated data to several models that used either areal grid cells or 1-km route segments (to approximate points) as sample units (Table 1). The Crosby and Porter (2018) model explicitly separates the probability of track-laying δ from the probability of occupancy (or use), and models the former as a function of the number of days elapsed since the last snowfall event, assuming that track-laying accumulates, and tracks persist over time. Because tracks are laid over time, occasions do not represent points in time; therefore, we did not impose any spatio-temporal exclusion among packs. Surveys occurred 1–3 days after a snowfall event, and the probability δ_j of observable tracks being laid was a function of time elapsed since the last snowfall event, and was also a Markovian process where track-laying was more likely at segment i if tracks were also laid at segment $i - 1$ (Crosby and Porter 2018). For the purpose of simulating data, we defined ψ_i and θ_{ij} as in Crosby and Porter (2018); ψ_i was the probability that a transect passed through ≥ 1 home range, and $\theta_{ij} = P(u_{ij} = 1)$ was the probability that route segment j on route i intersected ≥ 1 home range. In our simulated data, $u_{ij} = 1$ when a segment j in sample unit i intersected a home range. The model to fit the data specified spatial autocorrelation in θ_{ij} as a conditional autoregressive (CAR) process

Table 1. Sampling design for models fitted to simulated occupancy data. All simulated data were collected along routes based on the model of Crosby and Porter (2018) and were translated to grid-level data where necessary. Grid cells were either 100, 170, or 350 km² and routes were 40 km long. Grid sub-cells were either 4 or 9 sub-divisions of the grid cell sample units.

Model	Sample unit	Sub-unit	Description and assumptions
1	Grid cell	None	Simple occupancy model.
2	Route segment	None	Simple occupancy model.
3	Grid cell	None	Spatial conditional autoregressive (CAR) process on occupancy probability (ψ_i), with adjacent cells as neighbors.
4	Route segment	None	Spatial CAR process on ψ_i , with adjacent route segments as neighbors.
5	Grid cell	Grid sub-cell	Spatial replicates (sub-units) nested in temporal replicates. Occupancy (use) status can change, with use probability (θ), but it is assumed that if a cell is occupied, then all spatial subunits are occupied.
6	Route	Route segment	Same assumptions as model 5.
7	Grid cell	Grid sub-cell	Temporal replicates nested in spatial replicates. Occupancy status (z_i) cannot change, but only some subunits are assumed to be occupied, with subunit use probability θ_{ij} .
8	Grid cell	Route segment	Same assumptions as model 7.
9	Route	Route segment	Same assumptions as model 7.
10	Grid cell	Route segment	Identical to model 8 but with spatial CAR process on subunit use probability (θ_{ij}).
11	Route	Route segment	Identical to model 9 but with spatial CAR process on θ_{ij} .

where $\text{logit}(\theta_{ij}) = \alpha_0 + \rho_{ij}$, where ρ_{ij} was given an intrinsic CAR prior, with weights equal to 1 for neighbors and 0 otherwise (Crosby and Porter 2018). Detection probability was high ($p = 0.85$), given $u_{ij} = 1$. To investigate the effects of increased effort, we simulated either 30 or 50 observation routes, each 40 km long, and either 3 or 6 survey occasions. We generated 100 data sets for each simulated scenario, and fitted the data to 11 different models, with either grid-based or route segment-based sample units (Table 1). We carried out simulations in program R (R Core Team 2018) and fitted models using the R package NIMBLE (de Valpine et al. 2019).

We simulated scenarios for a 17,000-km² study area where the square sampling grid size was either 100, 170, or 350 km², the number of simulated home ranges was either 20, 40, or 80, and home range placement was clustered based on a simulated covariate whose value was moderately spatially correlated. We generated home range size to approximate empirical observation from a gamma distribution with mean = 170 km² and variance = 2,890. Mean pack size was 4, and we use an equal probability multinomial distribution to distribute the wolves among the modeled packs. Thus, simulated population sizes were exactly 80, 160, or 320 individuals. We estimated abundance as $N = \bar{x}\theta\sum_{i=1}^I \psi_i A_i / \bar{h}$, where I is the number of sample subunits. We expected the most severe positive bias in cases where sample units were large and where abundance was low (because potential bias in occupancy estimate is greatest under these conditions) and when home ranges were small (because the consequences of positive bias in occupancy estimates are most consequential when more home ranges can be packed into the landscape).

Application 1: Wisconsin Wolf Management Zones 1 and 3

We conducted an analysis of roadside snow-tracking data collected during winter 2016–2017 in Wolf Management Zones (WMZ) 1 and 3 in Wisconsin (Fig. 1). Snow tracking occurred in blocks conveniently defined by rivers and roads, but effort recorded at the tracking block scale is not easily translated to consistent sample grids unless the effort is spatially referenced. We digitized survey routes in WMZ 1 and 3 for 2016–2017 so that we could define effort (km traveled) at any scale or use survey routes themselves as sample units. We spatially referenced observations at locations where wolves entered, exited, or crossed roadways. Multiple surveys in tracking blocks usually were separated by ≥ 7 days. We defined grid cells independently from tracking blocks, and survey periods as the first 2 weeks (day 1–day 14) and last 2 weeks (day 15–end of month) of each month from 15 Nov 2016–31 Mar 2017, for 9 occasions. Consequently, the final occasion in each month was marginally longer than the first, and a survey occasion in a grid cell could conceivably include >1 survey.

We fitted each of the 3 occupancy models (above), without covariates, to the snow tracking data. We modeled detection probability as a function of effort, where effort was

the log-transformed total length of survey routes in a cell. To scale occupancy estimates up to abundance, we used $\mu_{HR} \sim N(172.5, 17.7)$, based on global positioning system observations of 36 packs from 2015–2017, and $\mu_{packsize} \sim N(3.9, 0.106)$, based on pack size inferred from tracks observed from 230 putative packs in the winter of 2016–2017 (Wiedenhoef et al. 2018). Territory overlap was not known, so we did not adjust estimates to account for possible pack overlap as we did in the simulations.

We also fitted models 1, 7, and 8 (Table 1) to the above snow tracking data. We included road density and forest cover covariates to predict ψ_i , and modeled detection probability as a function of effort (for areal sample units). For models 7 and 8, no effort covariate was required because sample sub-units were all 1 km in length. To scale occupancy estimates up to abundance, we used prior distributions of $\mu_{HR} \sim \text{Gamma}(100, 100/172.5)$ and $\mu_{packsize} \sim \text{Gamma}(1500, 1500/3.9)$; these are almost identical to the normal distributions previously specified.

Application 2: Wisconsin Range-Wide Abundance Estimate

We also used snow-tracking data from winter 2018–2019 to estimate wolf abundance for the entire range of wolves in Wisconsin, where winter snow-tracking occurred (Fig. 1). Tracking occurred as described above, and we directly recorded or digitized tracking survey routes *post hoc* so that we could quantify survey effort. We used a 100-km² hexagonal grid to delineate sample units and derived 2 landscape covariates (forest and agriculture) from National Land Cover Data (Yang et al. 2018). We fitted 9 static occupancy models (Table 2; model code available in Supporting Information) with a detection model that included a constant detection parameter, an effort covariate (log of tracked distance), or a sample unit by occasion random effect. We modeled occupancy probability ψ as constant, a function of landscape covariates and road density, or with a first-order spatial random effect. Occasions were 7-day periods from 1 Nov 2018 to 11 Apr 2019, which resulted in 24 sampling occasions. As described previously, we scaled occupancy up to abundance using $\mu_{HR} \sim N(172.5, 17.7)$

Table 2. Occupancy models fitted to winter 2018–2019 range-wide wolf tracking data in Wisconsin, USA. Landcover covariates were derived from 2016 National Landcover Data. In the detection model, effort = log(km surveyed), and ϵ_{ij} are site- and time-specific random effects. For the occupancy model, road = km of road/km², ag = agriculture and developed land cover, for = forest land cover, and CAR is a conditional autoregressive spatial process.

Model	Detection sub-model	Occupancy sub-model
1	effort + ϵ_{ij}	constant
1b	effort	constant
2	constant	constant
3	effort + ϵ_{ij}	for + ag + road
3b	effort	for + ag + road
4	effort + ϵ_{ij}	CAR
4b	effort	CAR
5	effort + ϵ_{ij}	for + ag + road + CAR
5b	effort	for + ag + road + CAR

and $\mu_{\text{packsize}} \sim N(3.9, 0.106)$. We compared the competing models using the Watanabe-Akaike Information Criterion (WAIC; Watanabe 2010, Hooten and Hobbs 2015). Wolf capture and collaring followed the Wisconsin Department of Natural Resources' protocol for safe and humane capture, handling, and sampling of gray wolves approved by the Department's Animal Care and Use Committee (GrayWolf_SOP_11).

RESULTS

Simulation 1: Areal Sample Units

Abundance was overestimated in most but not all scenarios, with estimates >3 times the true values in the most extreme case. Although reliability of abundance estimates was highly dependent on simulated scenarios, several clear patterns emerged (Fig. 2). First, given the same value for true abundance, larger sampling grid cells resulted in larger estimates of abundance. Second, regardless of true abundance, there was positive bias in estimates when home range size was small, and positive bias was exacerbated when sample units were large. Third, abundance was nearly always overestimated when true population size was small, but when population size was large, positive bias occurred only when home range size was small. Fourth, as expected, abundance estimates from the static occupancy model were always greater than estimates from the dynamic model. Also, estimates from the dynamic model generally were larger than estimates from the multi-scale model. Fifth, decreasing detection probability resulted in fewer detections across the spatial replicates and consequently resulted in lower abundance estimates and lower, but still substantial, positive bias (Fig. 3). Conversely, simulating greater movement rates (higher availability across the home range) resulted in larger abundance estimates and greater positive bias (Fig. 3). Restricting sampling to alternating grid cells increased uncertainty and only minimally reduced bias (Fig. S1, available in Supporting Information).

When packs were not cohesive (availability of individuals was independent of other pack members), there were more detections and consequently greater precision, but also greater bias (Fig. S2, available in Supporting Information). For availability model 3 (simulating equal availability within all subunits of occupied cells), the static occupancy model clearly estimated cumulative occupancy, whereas absolute bias was minimal (<0.1) for the dynamic and multi-scale models (Fig. S2). Bias was smaller for the dynamic model, but precision was better for the multi-scale model (Fig. S2).

Simulation 1: Linear Sample Units

When we simulated observation data from road segments, a hierarchical sampling design where route segments served as sub-samples approximating point samples, in some cases (models 8 and 9; Table 1) greatly mitigated the extreme positive bias evident when sample units were strictly areal. Model 8 performed best, where route segments were secondary sample units embedded in grid-cell primary sample units. Under this scenario, mean relative bias in posterior means was <0.1 in most cases, except when population size

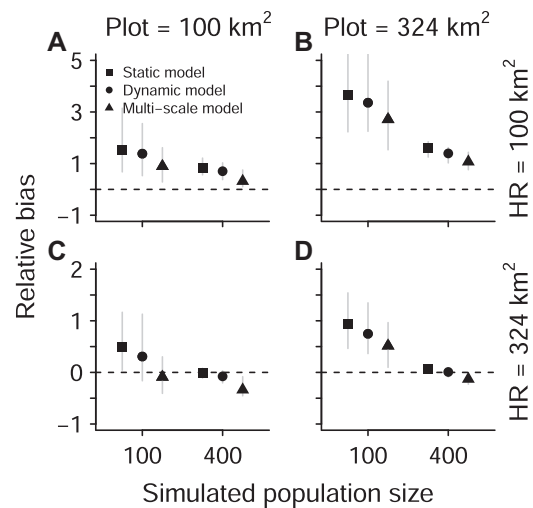


Figure 2. Relative bias in estimated abundance derived from 3 different occupancy models for various combinations of simulated small (A, C) or large (B, D) sample units (plot), small (A, B) or large (C, D) home ranges (HR), and sparse and dense populations, for simulated pack-dwelling wolves. Dashed line denotes zero bias, and error bars the range of bias observed across all simulations.

and home range size was small and sampling unit size was large (Fig. 4). Model performance was somewhat poorer when we modeled route segment use with a spatial CAR process than when use probability was not dependent on adjacent segment use. Model 10 (Table 1) was identical to model 8, except that segment use included a spatial CAR model. There was greater positive bias from model 10 than from model 8, when population size was small, and greater negative bias when population size was large, and home range size was more influential (Fig. S3, available in Supporting Information).

Previously we showed that when the closure assumption was severely violated, greater sampling effort resulted in greater positive bias. The same was true for many of the scenarios considered in this section but not universally so. In particular, bias did not increase with sampling effort for designs where route segments were the spatial replicates, even when the primary units were grid cells (Fig. S4, available in Supporting Information).

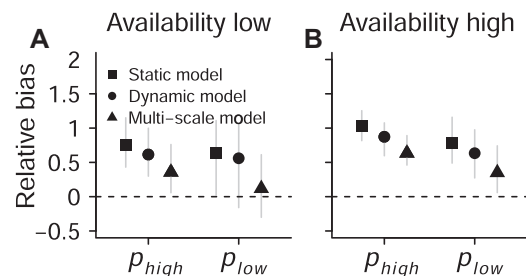


Figure 3. Relative bias in estimated abundance derived from 3 occupancy models when detection probability (p) was high or low, and individuals within simulated packs were moderately (A) or highly (B) mobile within their home range. In each simulated scenario, true home range size and sample grid size was 225 km^2 and true abundance was 200 individuals in packs with mean size = 3.8 individuals. Dashed line denotes zero bias, and error bars the range of bias observed across all simulations.

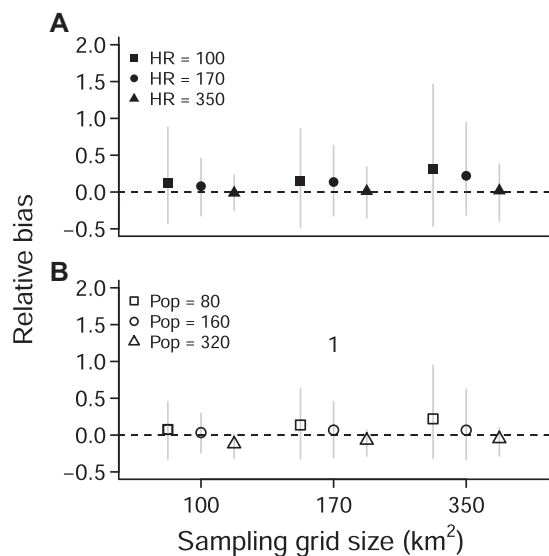


Figure 4. Relative mean bias in estimated population size for the simulation scenario where parameters were estimated from a multi-scale occupancy model where sample units were linear road segments nested in areal grid cells. In A, simulated population size (pop) was small (80 individuals in 20 packs). In B, mean territory size (HR) size was 170 km² in all scenarios. Dashed line denotes zero bias, and error bars the range of bias observed across all simulations.

Application 1: Wisconsin Wolf Management Units 1 and 3

As was the case for simulation results, estimates of wolf population size in WMZ 1 and 3 increased when the unit size of the sampling grid increased, although only slightly (Fig. 5; Table S1, available in Supporting Information). Estimates from the dynamic model also were slightly smaller than for the static occupancy model, as expected, and estimates from the multi-scale model were much smaller than for either of the other 2 models. For the static and dynamic models, 95% credible intervals included the minimum count of 499–512 for WMZ 1 and 3 combined, regardless of the sampling grid size (Fig. 5). Estimates from the multi-scale model were substantially lower than the minimum count, and the minimum count was included in the 95% credible interval (CrI) only when the sampling grid cells were large. Also, convergence ($\hat{R} < 1.1$) was more difficult to achieve with the multi-scale model than with the other 2 models. When we fitted data from only a subset of grid cells to the models, abundance estimates generally were marginally reduced compared to estimates from the full dataset (Fig. 5), but posterior means varied by as much as 20%, depending on which specific grid cells were selected.

For models fitted to wolf tracking data from WMZ 1 and 3, estimated population size was largest for the standard occupancy model, and estimates increased with increasing sample unit size (Fig. 6). For the multi-scale model with areal sub-units, the estimate of population size also increased with grid cell size, and the estimate was lower when there were 9 as opposed to 4 spatial sub-units. The model with 1-km road segments representing spatial subunits (to approximate points sample units), estimated the lowest

overall population size (unrealistically low) but was least sensitive to the size of the primary sample unit.

Application 2: Wisconsin Range-Wide Abundance Estimate

For wolf data from winter 2018–2019, the model with the lowest WAIC value included an effort effect on detection probability and landscape covariate effects and a spatial CAR effect on occupancy probability (model 5b in Table 2). The correlation parameter of the spatial CAR effect did not adequately converge, however, and thus we present results from the second best-supported model (model 3b in Table 2), which dropped the CAR effect. Detection probability was positively related to effort ($\alpha_{\log\text{effort}} = 0.85$; 95% CrI = 0.72–0.99). Occupancy probability was positively related to forest cover ($\beta_{\text{forest}} = 0.93$; 95% CrI = 0.47–1.45) and negatively related to agricultural and developed land ($\beta_{\text{ag}} = -0.8$; 95% CrI = -1.19–0.41) and road density ($\beta_{\text{road}} = -1.13$; 95% CrI = -2.04–0.29). The abundance estimate (mode; 95% CrI) was 1,047 (835–1,333). The range-wide count for winter 2018–2019 was 914–978 wolves (Wiedenhoef et al. 2019).

DISCUSSION

Occupancy estimates are used as alternatives to abundance estimates because they obviate the need to obtain individual identifications, and can effectively integrate various data types into a single analysis framework (MacKenzie et al. 2006, Pacifici et al. 2017). Interpreting surrogates or indices as accurate estimates of abundance warrants caution (MacFarland and Van Deelen 2011, Steenweg et al. 2018), and our results suggest that the same is true when explicitly scaling occupancy estimates

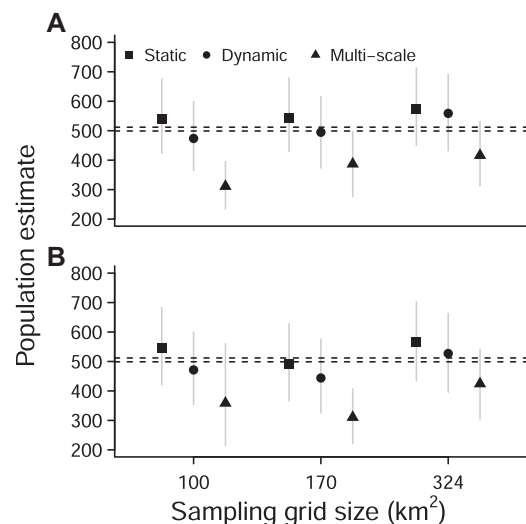


Figure 5. Abundance estimates for wolves in Wolf Management Zones (WMZ) 1 and 3 in northern Wisconsin, USA, winter 2016–2017, scaled up from 3 different occupancy models, when sampling occurs in a complete grid (A) or only in alternating grid cells (B). Error bars represent 95% highest probability density credible intervals, and dashed lines represent lower and upper limits of the summed minimum count, based on territory mapping, for WMZ 1 and 3.

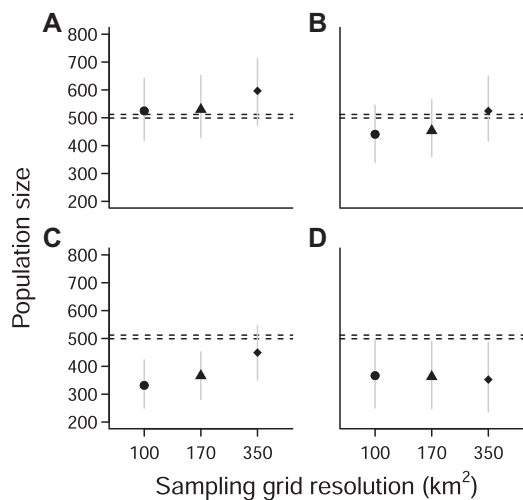


Figure 6. Occupancy abundance estimates for wolves in Wolf Management Zones (WMZ) 1 and 3 in northern Wisconsin, USA, winter 2016–2017. The estimates are from a simple static model (A); a multi-scale model where we modeled use at 4 (B) or 9 (C) sub-divisions of each primary sampling grid cell; and a multi-scale model where sub-units were 1-km survey route segments, to approximate point sample units (D). Error bars represent 95% credible intervals, and the 2 dashed lines represent the lower (499) and upper (512) bounds of the minimum count for WMZ 1 and 3.

up to estimate population abundance. MacKenzie et al. (2006:42) stated that when territorial animals are surveyed, and size of the sample units matches territory size, then the number of occupied sites is “virtually equivalent to estimates of the number of territorial animals or pairs,” and this assumption underlies occupancy-based estimates of abundance (Rich et al. 2013, Ausband et al. 2014). Under reasonable scenarios of possible movement within territorial home ranges, our results suggest that the assumption that the number of occupied areal sample units equates the number of occupied territories holds only under certain relatively narrow conditions, and those conditions are not entirely controllable. Instead, reliability of occupancy-abundance estimates is sensitive to mean home range size, arrangement of home ranges, sampling grid cell size, and actual abundance. Although we primarily discuss scaling occupancy to abundance of individuals, an intermediate step is scaling occupancy to abundance of packs. Concerns about bias apply equally to both cases.

The fundamental reason for positive bias in occupancy estimates is violation of the closure assumption. Movements of individuals among sample units, especially when sample units constitute a continuous grid, leads to estimation of cumulative use rather than instantaneous occupancy (Efford and Dawson 2012). Cumulative probability of occupancy (or the proportion of the area used) is not the problem. It should be the correct metric, and the claim by MacKenzie et al. (2006:24) that the number of occupied cells equals the number of home ranges of territorial animals should be approximately correct, if cell size equals mean home range size, and crucially, if home ranges are arranged with no

interstitial spaces. When home ranges are dispersed with substantial interstitial spaces (e.g., when population density is low), then detections within a single home range can occur in >1 adjacent cell, and what is estimated is the number of sample units that contain any portion of a home range (Fig. 7), which is greater than the number of cells that contain a home range center. Conceivably, a pack whose home range is centered on the boundary of 4 sample units could be detected in each of adjacent units, and resulting estimated area occupied could be 4 times the home range size. The positive bias in the occupancy estimate is exacerbated when those sample units are excessively large. Bias is also exacerbated when detection probability is high, because a single pack is more likely to be detected in multiple sample units. Restricting the number of sample units so they are non-adjacent only marginally mitigated the bias because movement in and out of sample units results in negative bias in detection probability, or, for designs with spatial replication, incomplete availability across spatial replicates results in the same issue (Kendall and White 2009).

Use of spatial replicates to collect robust design data and estimate abundance from ψ_0 or ψ_θ also did not adequately mitigate bias, in most cases. In principle, ψ_0 should estimate instantaneous occupancy, and ψ_θ should exclude the unused portion of primary sample units from the estimated area occupied. Indeed, these estimates resulted in minimal bias in the scenario where individuals were available for detection in all subunits of the sample unit they occupied at a given occasion, and not in subunits of adjacent cells. Such a scenario is largely unrealistic when sample units form a comprehensive grid. Home ranges will often straddle the boundaries of multiple adjacent sample units, with consequent positive bias almost inevitable (Hayes and Monfils 2015), unless the landscape is largely saturated.

The consequences of closure assumption violations depended on home range size, population size and dispersion, and sample unit size, only the latter of which is controllable. Positive bias in occupancy estimates is more consequential for abundance estimates when home ranges are small because more small than large home ranges fit into the estimated area occupied. Smaller home ranges do not preclude an occupancy approach to abundance estimation, but that situation allows less room for error in selection of sample units.

Association of increasing positive bias with larger sample unit sizes and smaller home range sizes seems at odds with findings of previous investigations (Efford and Dawson 2012, Hayes and Monfils 2015) where occupancy estimates increased as the ratio of home range size to sample unit size increased. Efford and Dawson (2012) did not investigate scenarios with comprehensive sample unit coverage (e.g., a continuous grid), as in our scenarios. Hayes and Monfils (2015) concluded increasing cumulative occupancy as home range size increased, but their definition of positive bias depended on their definition of true occupancy, apparently instantaneous occupancy, which would not change with home range size. Logically, when sample unit size is much larger than territory size, a single group or

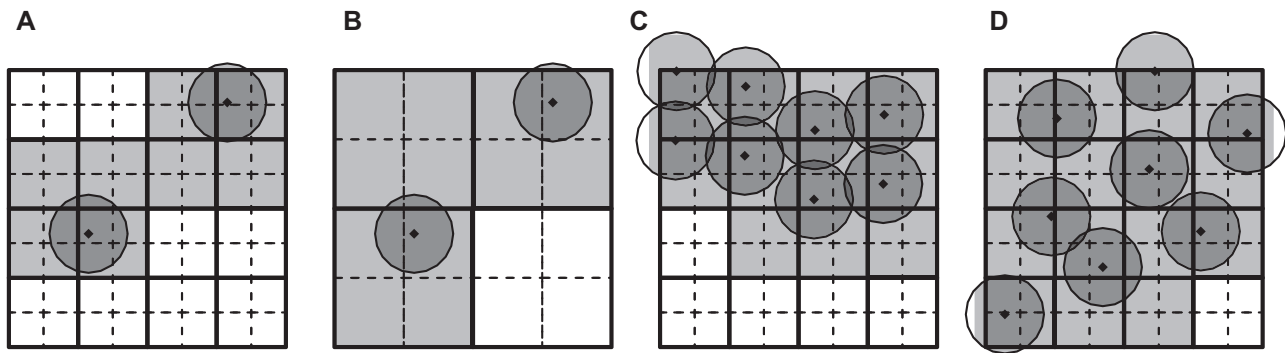


Figure 7. Overlap of hypothetical home ranges with multiple sample units. The solid grid lines depict adjacent sample units, dashed lines depict spatial replicates, circles depict home ranges with the same area as the grid cells, and points represent home range centers. In A, 2 home ranges overlap 8 grid cells, thus occupancy is 0.5, and the estimated number of territories (based on the number of occupied cells) is 4 times the true value. In B, the sample cell size is quadrupled, but the number of home ranges is unchanged. In this case occupancy is 0.75 and the estimated number of home ranges is 12, or 6 times the true value. In C, the cell size is the same as in A, but the number of home ranges is increased, so that that it is more likely that adjacent cells actually contain a home range center. In this case occupancy is 0.69, and bias in the estimated number of home ranges is reduced to 11/8, or 1.38 times the true value. In D, there are again 8 home ranges, but they are less aggregated than in C. In this case occupancy is 0.94, and the number of estimated home ranges is 1.9 times the true value.

individual occupying a sample unit should lead to inflated estimates of the amount of occupied landscape (Fig. 7). In the extreme, a single sample unit that includes the entire study area is occupied with probability 1 if a single individual is detected. In that case, the estimate of population size is completely determined by the number of territories that can fit in the landscape, and the mean group size. The smaller number of large home ranges that can fit in the landscape thus imposes an upper limit on positive bias that is much lower than the limit imposed by small home ranges, even when sample unit size matches home range size in both cases. Of course, increasing home range overlap increases the imposed upper limit.

Given the results of our findings of positive bias in many scenarios, the logical consistency of observed positive bias in occupancy and abundance estimates when sample units are continuous, and the various ways that different home-range and sample unit configurations can result in identical occupancy estimates (Efford and Dawson 2012), it seems counter-intuitive that scaling occupancy to estimate abundance of wolves in Wisconsin and Montana produces reasonable and plausible estimates (this study, Rich et al. 2013, Montana Fish, Wildlife, and Parks 2018). Nevertheless, there are several reasons why these real-world applications work. In the simulations, positive bias was greatest when true population size was small and sample grid cell size was larger than home range size, and when home range size was small, regardless of grid cell size. When population size is small, closure assumption violations can result in occupancy of multiple adjacent cells by a single group or individual (Fig. 7). The same is true at high density, but, crucially, when territories nearly saturate the landscape, closure assumption violations are less consequential, and tend to offset each other for adjacent sample units. In both Wisconsin and elsewhere, it is likely that wolf home ranges are aggregated in the highest quality habitat where mortality risk is lowest, to the extent that these areas are nearly saturated (Mladenoff 2009, O'Neil et al. 2017). Montana uses

a large (600 km²) sampling grid resolution, multiple state-wide observation data sources, and informative landscape covariates to predict occupancy of wolves and scale up to abundance (Rich et al. 2013, Montana Fish, Wildlife, and Parks 2018), and most occupied sample units are aggregated in the western part of the state (Rich et al. 2013: figure 1). Our analysis of range-wide wolf abundance in Wisconsin for winter 2018–2019 also suggested that informative covariates for ψ mitigated potential positive bias, and that the northern Wisconsin landscape is nearly saturated with wolf territories. Consequently, our abundance estimate was reasonably close to the minimum count derived for that winter. Saturation results in a distribution with minimal interstitial spaces (Fig. 7C), where closure violations are of minimal consequence. Landscape features influence distribution, with 2 consequences that both mitigate potential positive bias. First, wolf detections are most likely in sample units with favorable landscapes, so closure violation is less likely. Second, landscape covariates lead to better prediction in sample units without detections.

Because the abundance estimate (of either packs or individuals) is a *post hoc* adjustment of an occupancy estimate, the primary source of potential bias is the occupancy estimate. Scaling relies crucially on accurate estimates of home range features, however, and those features can exacerbate bias, in some cases. In our simulations we considered a relatively narrow range of non-overlapping home range sizes, reasonable for wolves in Wisconsin and the Great Lakes region (Wiedenhoef et al. 2018), and we considered only circular home ranges. We expect that the same general patterns apply for larger home ranges, and larger sample unit sizes. For example, there is not evident extreme bias in the Montana estimates (Montana Fish, Wildlife, and Parks 2018), where sample units and home ranges are nearly twice the size of the largest we considered here. In reality, home ranges may be highly variable in size and shape, and implications for occupancy abundance estimates are not entirely clear. Although additional simulations would be

useful, greater variability in home range sizes should not generate additional bias but might increase uncertainty in estimation of mean home range size that propagates into the abundance estimate. If home ranges sizes vary substantially, either spatially or based on habitat features, it may be necessary to incorporate variation in the home range model more rigorously. Similarly, we see no reason that irregular home range shapes should greatly influence bias, although we suggest that when home ranges are highly irregular, the number of home ranges that fit into the landscape may be decreased. For example, a home range described by a maximum convex polygon may contain area that is not used by a pack, but is effectively co-opted from use by another pack. We suggest that in cases with highly irregular home range shapes, the mean home range size should be calculated including such unused area, rather than strictly using a home range model such as a kernel density estimator. We also simulated a relatively small range of abundances, with little or no home range overlap. We do not think that the absolute size of the population influences bias as much as the distribution of the population; a situation with highly dispersed packs and extensive interstitial space is more subject to positive bias than a situation where packs are clustered. Overlapping home ranges probably do not influence occupancy estimation itself (other than the fact that overlapping home ranges take up less space than dispersed home ranges) but obviously have implications for scaling occupancy to abundance.

Although real-world applications produced plausible answers (this study, Rich et al. 2013, Montana Fish, Wildlife, and Parks 2018), it is worth reiterating that in simulations, positive bias was greatest when true abundance was small. Recovery plans for species listed under the Endangered Species Act may include specific numeric goals, and positive bias in the estimators could very easily lead to erroneous and consequential conclusions about species recovery, when the population is small and home ranges are widely scattered. On the other hand, it is encouraging that rigorous density estimates from a spatial capture-recapture model, which are not subject to closure assumption violations in the same way that non-spatial models are (Royle et al. 2014), and estimates from an occupancy model for a low-density carnivore resulted in very similar management recommendations (Linden et al. 2017).

In many estimation problems, incorporating multiple sources or multiple types of data can improve precision of estimates. For example, data on individual identities or telemetry locations for some individuals improve estimates from spatial capture-recapture or spatial count models (Chandler and Royle 2013, Sollmann et al. 2013), multiple sources of occupancy data expand the scope and precision of occupancy models (Pacifi ci et al. 2017), and integration of multiple data types into population models improves estimates or permits estimates of otherwise inestimable parameters (Schaub and Abadi 2011, Horne et al. 2019). Although implementation is beyond our present scope, we briefly mention 2 data sources that possibly could be incorporated in our approach, either to improve precision or

to reduce bias. The first is observations from motion-sensitive cameras, which should be relatively straightforward to incorporate by modifying the observation model. The second is the number of wolves detected at each set of observed tracks, which seems less straightforward. Perhaps informally, such counts could mark or partially mark packs and identify when packs cross over sample unit boundaries, thereby allowing data thinning to possibly mitigate closure violations. More formally, marked tracks might be fitted in a joint occupancy capture-recapture model, where the capture-recapture model estimates the number of packs, and the occupancy model either estimates occurrence probability (our current model) or pack abundance, if modeled as a Royle-Nichols model (Royle and Nichols 2003). We are skeptical of such a joint approach, for 2 reasons. First, it is unlikely that the number of tracks reliably and objectively identifies individual packs. Second, the key assumptions of the Royle-Nichols model (detection varies with abundance, independent detections) are very likely to be violated with wolf tracking data.

Given the real potential for positive bias in abundance estimates described here, outside of a narrow set of circumstance, it seems advisable to collect information about individual identifications if possible, to estimate abundance directly. Because such data are typically more expensive to collect than detection–non-detection data, there is strong motivation to find ways to use the less expensive data for management decisions. We suggest considerable caution when scaling occupancy estimates to abundance based on mean territory size and mean groups size. In the case of Wisconsin and Montana, other abundance estimates were available, where the precise extent of bias is unknown but probably not extreme (i.e., it is unlikely that index counts are 2 times the true population size). Thus, there is a basis for judging, in those specific cases, that abundance scaled up from occupancy estimates are highly unlikely to suffer from the extreme positive bias demonstrated for some simulation scenarios. Montana and Wisconsin also plausibly meet the narrow set of circumstances where occupancy reliably scales to abundance. Of course, scaling occupancy up to abundance also relies crucially on accurate estimates of mean home range size and mean pack size, which may be density-dependent, and are not generally obtainable from detection–non-detection data.

Alternatively, in some cases it may be reasonable to abandon management based on numeric goals that require abundance estimation, and instead focus on distribution goals. Given that abundance and occupancy estimates can result in similar management recommendations (Linden et al. 2017), and given the reduced effort and expense of collecting detection–non-detection data as opposed to capture-recapture data, it may suffice to develop goals based on where occurrence is desirable. Management could then target actions to maximize occupancy probability in preferred areas.

Managing for occupancy or distribution may be logistically more feasible than making potentially problematic assumptions to estimate abundance from occurrence data, but

it may not be a politically or socially feasible goal for a species like wolves. Wolves in the Upper Great Lakes have been listed under the Endangered Species Act as federally endangered for most of the last several decades (Refsnider 2009, U.S. Fish and Wildlife Service 2019), and federal delisting criteria and state management plans include specific numeric population targets (Wisconsin Department of Natural Resources 1999, U.S. Fish and Wildlife Service 2019). Ensuring that delisting criteria or numeric management goals are met requires that either a population distribution goal is framed in such a way as to guarantee that meeting the distribution goal also meets the population numeric goal, or that population size is directly estimated. Territory mapping from snow-tracking and ancillary information is one approach to abundance estimation (Wydeven et al. 2009), and scaling occupancy to abundance is an alternative approach that can provide plausible estimates under certain circumstances (this study, Rich et al. 2013, Montana Fish, Wildlife, and Parks 2018). **The former approach requires extensive tracking, enumeration of wolves from tracks, and somewhat subjective assignment of all tracks into presumed pack territories. The latter approach requires less tracking effort than territory mapping and does not require subjective pack assignments, but it does require independent estimates of mean pack and home range size, and potentially is subject to positive bias with small and scattered populations.** Territory mapping may be an appropriate method for small and scattered populations, but where reasonable and justifiable (as in our Wisconsin example) the scaled occupancy approach is preferable because it avoids the potential subjectivity of territory assignments and provides a rigorous and honest estimate of uncertainty.

MANAGEMENT IMPLICATIONS

Occupancy models are a conceptually attractive, and potentially cost-effective alternative to capture-mark-recapture methods for providing information about abundance or distribution. If distribution estimates are adequate for making management decisions, then occupancy does appear to provide a cost-effective and accurate framework. But if management decisions require knowledge about abundance, then occupancy models provide an acceptable framework only under certain restricted circumstances when density is sufficient to nearly saturate the landscape or when landscape covariates effectively aggregate the population and accurately predict occupancy. In particular, a scaled occupancy approach will be prone to positive bias for a scattered, low-density, highly mobile population.

ACKNOWLEDGMENTS

J. E. Wiedenhoft curated and coordinated collection of wolf tracking data. J. L. Price Tack and 2 anonymous reviewers provided comments and suggestions that substantially improved the manuscript. This work was funded by the United States Fish and Wildlife Service Wildlife Restoration Program.

LITERATURE CITED

- Adams, M. J., D. A. W. Miller, E. Muths, P. S. Corn, E. H. C. Grant, L. L. Bailey, G. M. Fellers, R. N. Fisher, W. J. Sadinski, H. Waddle, and S. C. Walls. 2013. Trends in amphibian occupancy in the United States. *PLOS ONE* 8:e64347.
- Ausband, D. E., L. N. Rich, E. M. Glenn, M. S. Mitchell, P. Zager, D. A. W. Miller, L. P. Waits, B. B. Ackerman, and C. M. Mack. 2014. Monitoring gray wolf populations using multiple survey methods. *Journal of Wildlife Management* 78:335–346.
- Chandler, R. B., and J. A. Royle. 2013. Spatially explicit models for inference about density in unmarked or partially marked populations. *Annals of Applied Statistics* 7:936–954.
- Chen, G., M. Kéry, M. Plattner, K. Ma, and B. Gardner. 2013. Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology* 101:183–191.
- Crosby, A. D., and W. F. Porter. 2018. A spatially explicit, multi-scale occupancy model for large-scale population monitoring. *Journal of Wildlife Management* 82:1300–1310.
- de Valpine, P., D. Turek, C. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2019. Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics* 26:403–413.
- Denwood, M. J. 2016. runjags: an R package providing interface utilities, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software* 71. <<http://runjags.sourceforge.net>>
- Durso, A. M., J. D. Willson, and C. T. Winne. 2011. Needles in haystacks: estimating detection probability and occupancy of rare and cryptic snakes. *Biological Conservation* 144:1508–1515.
- Efford, M. G., and D. K. Dawson. 2012. Occupancy in continuous habitat. *Ecosphere* 3:1–15.
- Falke, J. A., L. L. Bailey, K. D. Fausch, and K. R. Bestgen. 2012. Colonization and extinction in dynamic habitats: an occupancy approach for a Great Plains stream fish assemblage. *Ecology* 93:858–867.
- Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation* 116:195–203.
- Hayes, D. B., and M. J. Monfils. 2015. Occupancy modeling of bird point counts: implications of mobile animals. *Journal of Wildlife Management* 79:1361–1368.
- Hooten, M. B., and N. T. Hobbs. 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs* 85:3–28.
- Horne, J. S., D. E. Ausband, M. A. Hurley, J. Struthers, J. E. Berg, and K. Groth. 2019. Integrated population model to improve knowledge and management of Idaho wolves. *Journal of Wildlife Management* 83:32–42.
- Kendall, W. L., and G. C. White. 2009. A cautionary note on substituting spatial subunits for repeated temporal sampling in studies of site occupancy. *Journal of Applied Ecology* 46:1182–1188.
- Latham, M. C., A. D. M. Latham, N. F. Webb, N. A. McCutchen, and S. Boutin. 2014. Can occupancy-abundance models be used to monitor wolf abundance? *PLOS ONE* 9:e102982.
- Latif, Q. S., M. M. Ellis, and C. L. Amundson. 2016. A broader definition of occupancy: comment on Hayes and Monfils. *Journal of Wildlife Management* 80:192–194.
- Linden, D. W., A. K. Fuller, J. A. Royle, and M. P. Hare. 2017. Examining the occupancy-density relationship for a low-density carnivore. *Journal of Applied Ecology* 54:2043–2052.
- MacFarland, D. M., and T. R. Van Deelen. 2011. Using simulation to explore the functional relationships of terrestrial carnivore population indices. *Ecological Modelling* 222:2761–2769.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* 84:2200–2207.
- MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Academic Press, San Diego, California, USA.
- Mladenoff, D. J. 2009. Change in occupied wolf habitat in the Northern Great Lakes region. Pages 119–138 in A. Wydeven, T. R. van Deelen,

- and E. Heske, editors. Recovery of gray wolves in the Great Lakes region of the United States; an endangered species success story. Springer, New York, New York, USA.
- Montana Fish, Wildlife, and Parks. 2018. Montana gray wolf conservation and management—2017 Annual report. Montana Fish, Wildlife, and Parks, Helena, USA.
- Mordecai, R. S., B. J. Mattsson, C. J. Tzilkowski, and R. J. Cooper. 2011. Addressing challenges when studying mobile or episodic species: hierarchical Bayes estimation of occupancy and use. *Journal of Applied Ecology* 48:56–66.
- O'Neil, S. T., J. K. Bump, and D. E. Beyer. 2017. Spatially varying density dependence drives a shifting mosaic of survival in a recovering apex predator (*Canis lupus*). *Ecology and Evolution* 7:9518–9530.
- Pacifici, K., B. J. Reich, D. A. W. Miller, B. Gardner, G. Stauffer, S. Singh, A. McKerrrow, and J. A. Collazo. 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology* 98:840–850.
- Pavlack, D. C., J. A. Blakesley, G. C. White, D. J. Hanni, and P. M. Lukacs. 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. *Journal of Wildlife Management* 76:154–162.
- R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Refsnider, R. L. 2009. The role of the Endangered Species Act in Midwest wolf recovery. Pages 49–64 in A. Wydeven, T. R. van Deelen, and E. Heske, editors. Recovery of gray wolves in the Great Lakes region of the United States; an endangered species success story. Springer, New York, New York, USA.
- Rich, L. N., D. A. W. Miller, H. S. Robinson, J. W. McNutt, and M. J. Kelly. 2016. Using camera trapping and hierarchical occupancy modeling to evaluate the spatial ecology of an African mammal community. *Journal of Applied Ecology* 53:1225–1235.
- Rich, L. N., R. E. Russell, E. M. Glenn, M. S. Mitchell, J. A. Gude, K. M. Podrutzny, C. A. Sime, K. Laudon, D. E. Ausband, and J. D. Nichols. 2013. Estimating occupancy and predicting numbers of gray wolf packs in Montana using hunter surveys. *Journal of Wildlife Management* 77:1280–1289.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84:777–790.
- Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner. 2014. Spatial capture-recapture. Academic Press, Waltham, Massachusetts, USA.
- Schaub, M., and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology* 152:227–237.
- Sollmann, R., B. Gardner, A. W. Parsons, J. J. Stocking, B. T. McClintock, T. R. Simons, K. H. Pollock, and A. F. O'Connell. 2013. A spatial mark-resight model augmented with telemetry data. *Ecology* 94:553–559.
- Steenweg, R., M. Hebblewhite, J. Whittington, P. Lukacs, and K. McKelvey. 2018. Sampling scales define occupancy and underlying occupancy–abundance relationships in animals. *Ecology* 99:172–183.
- Theil, R. P., W. Hall, E. Heilhecker, and A. P. Wydeven. 2009. An isolated gray wolf population in central Wisconsin. Pages 107–117 in A. Wydeven, T. R. van Deelen, and E. Heske, editors. Recovery of gray wolves in the Great Lakes region of the United States; an endangered species success story. Springer, New York, New York, USA.
- Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications* 13:1790–1801.
- U.S. Fish and Wildlife Service. 2019. Endangered and threatened wildlife and plants; removing the gray wolf (*Canis lupus*) from the list of endangered and threatened wildlife. *Federal Register* 84:9648–9687.
- Watanabe, S. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* 11:3571–3594.
- Whittington, J., K. Heuer, B. Hunt, M. Hebblewhite, and P. M. Lukacs. 2015. Estimating occupancy using spatially and temporally replicated snow surveys. *Animal Conservation* 18:92–101.
- Wiedenhoef, J. E., S. Walter, N. Kluge, and M. Ericksen-Pilch. 2019. Wisconsin gray wolf monitoring report 15 April 2018 through 14 April 2019. Bureau of Wildlife Management, Wisconsin Department of Natural Resources, Madison, USA.
- Wiedenhoef, J. E., S. Walter, N. S. Libal, and M. Ericksen-Pilch. 2018. Wisconsin gray wolf monitoring report 15 April 2017 through 14 April 2018. Bureau of Wildlife Management, Wisconsin Department of Natural Resources, Madison, USA.
- Wisconsin Department of Natural Resources. 1999. Wisconsin wolf management plan. Wisconsin Department of Natural Resources, Madison, USA.
- Wydeven, A. P., J. E. Wiedenhoef, R. N. Schultz, R. P. Theil, R. L. Jurewicz, B. E. Kohn, and T. R. Van Deelen. 2009. History, population growth, and management of wolves in Wisconsin. Pages 87–105 in A. Wydeven, T. R. van Deelen, and E. Heske, editors. Recovery of gray wolves in the Great Lakes region of the United States; an endangered species success story. Springer, New York, New York, USA.
- Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian. 2018. A new generation of the United States National Land Cover Database: requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing* 146:108–123.
- Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution* 16:446–453.

Associate Editor: Ryan Long.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.