ECOSPHERE

Occupancy in continuous habitat

MURRAY G. EFFORD[1],† AND DEANNA K. DAWSON[2]

[1]Department of Zoology, University of Otago, P.O. Box 56, Dunedin 9054 New Zealand
[2]United States Geological Survey, Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, Maryland 20708 USA

Abstract.   The probability that a site has at least one individual of a species ('occupancy') has come to be widely used as a state variable for animal population monitoring. The available statistical theory for estimation when detection is imperfect applies particularly to habitat patches or islands, although it is also used for arbitrary plots in continuous habitat. The probability that such a plot is occupied depends on plot size and home-range characteristics (size, shape and dispersion) as well as population density. Plot size is critical to the definition of occupancy as a state variable, but clear advice on plot size is missing from the literature on the design of occupancy studies. We describe models for the effects of varying plot size and home-range size on expected occupancy. Temporal, spatial, and species variation in average home-range size is to be expected, but information on home ranges is difficult to retrieve from species presence/absence data collected in occupancy studies. The effect of variable home-range size is negligible when plots are very large (>100 × area of home range), but large plots pose practical problems. At the other extreme, sampling of 'point' plots with cameras or other passive detectors allows the true 'proportion of area occupied' to be estimated. However, this measure equally reflects home-range size and density, and is of doubtful value for population monitoring or cross-species comparisons. Plot size is ill-defined and variable in occupancy studies that detect animals at unknown distances, the commonest example being unlimited-radius point counts of song birds. We also find that plot size is ill-defined in recent treatments of "multi-scale" occupancy; the respective scales are better interpreted as temporal (instantaneous and asymptotic) rather than spatial. Occupancy is an inadequate metric for population monitoring when it is confounded with home-range size or detection distance.

† **E-mail:** murray.efford@otago.ac.nz

INTRODUCTION

Ecologists have long struggled with estimating the abundance of mobile or cryptic animals, and many have turned to measuring spatial distribution as a potentially robust and cheaper alternative. The use of distribution as a state variable has been enhanced by the development of elegant and flexible models that accommodate imperfect detection (MacKenzie et al. 2002, 2006, Tyre et al. 2003). These methods use repeated surveys at a sample of sites to estimate the true probability that a site is occupied by a species, the parameter known as 'occupancy'. Occupancy

is now widely seen as an alternative to population density (or abundance in a broad sense) in large-scale monitoring programs (MacKenzie and Nichols 2004, MacKenzie et al. 2005, Marsh and Trenham 2008, Morrison et al. 2008, Mattsson and Marshall 2009, Weir et al. 2009, O'Brien et al. 2010, Zylstra et al. 2010). While not exactly a population parameter, occupancy represents population status in a way that is potentially useful for management. Occupancy has an advantage over simple indices of population density if it is possible to separate confounded variation in detection probability from variation in the state variable.

Occupancy studies vary widely in their sampling designs, particularly in what they treat as a site. Among other possibilities, a site may be a discrete island or patch of habitat, a plot surrounded by continuous habitat, or a length of trail. Replication may be achieved by surveying repeatedly over time or simultaneously by multiple observers or methods; surveys of spatial subunits are also sometimes used as repeat samples. Repeated surveys yield a matrix of binary observations coding the detection or non-detection of a species at each site on each survey. Fitting the standard model (MacKenzie et al. 2002, Tyre et al. 2003) leads to estimates of occupancy (ψ) and the probability of detecting the species given that it is present at a site (p). Occupancy may vary between seasons or areas, perhaps in response to habitat covariates. A key aim of modeling is to separate potentially interesting variation in ψ from variation in the nuisance variable p.

Yoccoz et al. (2001) framed the problem of study design for biodiversity monitoring as three questions: Why monitor? What should be monitored? How should monitoring be carried out? The literature on the design of occupancy studies has focused on the "How" question, particularly on statistical power and the trade-off between number of sites and frequency of sampling (MacKenzie and Royle 2005, Bailey et al. 2007, Guillera-Arroita et al. 2010). Of equal or greater importance is the "What" question—specifically, the choice of a state variable to monitor, and the scale at which to measure it. To clearly define occupancy as a state variable, it is necessary to specify what is meant by a site, and whether occupancy is 'instantaneous' or integrated over time.

The two seminal papers on occupancy with imperfect detection (MacKenzie et al. 2002, Tyre et al. 2003) conceived the temporal dimension of occupancy differently. MacKenzie et al. (2002) were concerned primarily with the instantaneous occupancy of isolated patches. If movement of individuals across the site boundary changed the occupancy status of a site between surveys this was viewed as a violation of the 'closure' assumption. Tyre et al. (2003) were concerned with sites embedded in a matrix of similar habitat, and defined occupancy as a parameter that integrated over time. They observed "For species whose home ranges exceed the survey area [ψ_p] also includes the likelihood that the individuals are present within the survey area at the time it is conducted", thus anticipating later re-interpretations of the MacKenzie et al. (2002) approach in terms of robustness to temporary emigration (e.g., Kendall and White 2009).

The terminology of non-instantaneous occupancy has not been standardized. We use the term 'asymptotic' for this type of occupancy (i.e., the asymptote approached by cumulative observations of occupancy over time). The concept is analogous to the 'superpopulation' of animals intermittently available for capture in robust-design capture–recapture studies (Kendall 1999). Instantaneous occupancy is difficult to monitor in continuous habitat because estimation requires simultaneous or nearly simultaneous independent samples. Asymptotic occupancy is the parameter more commonly estimated and the one upon which we focus here.

Another aspect of the "What" question for occupancy studies in continuous habitat is plot size, about which little advice can be gleaned from the literature. MacKenzie and Royle (2005) stated that "[For] a species with relatively large home ranges compared with the size of the sampling units, the proportion of area 'used' over a longer timeframe may be close to 100% even though population size is very small". O'Connell and Bailey (2011) worried that plots smaller than a home range would violate the assumption of closure [i.e., constant occupancy] between replicate samples. Karanth et al. (2011) justified their choice of plot size in a landscape-scale study of tigers (*Panthera tigris*) thus "Because our goal was to measure true occupancy (proportion of area

occupied) rather than intensity of habitat use by tigers (MacKenzie and Royle 2005, MacKenzie et al. 2006), the cell size chosen was, on average, larger than the maximum expected tiger home range size." We know of no attempt to address the quantitative relationships between plot size, home-range size and occupancy.

The concept of plot size is itself somewhat complex. In some occupancy studies, plots are defined areas that are searched quickly for individuals of the target species present at a point in time (e.g., Tyre et al. 2003). This relies on direct observation of the target species, which may not be possible. Searching plots for sign, such as tracks, feces, or antler rubs, is an alternative; although the data resemble those from direct observation, they can raise issues of interpretation that we touch on in the Discussion. Other studies instead rely on passive detectors (traps, cameras, tracking tunnels, etc.) that record individuals of the target species when they visit a particular place. Plot size is then indeterminate and probably very small, barring long-range attraction of animals by means of lures. Plot size is also indeterminate when animals are detected at a distance and the distance threshold for detection is variable or unknown.

Our overall aim in this paper is to clarify the implications of home-range size and plot size for the design of occupancy studies in continuous habitat. The properties of the nominal state variable 'occupancy' differ greatly when plots are very large or very small relative to home-range size. Plot size is usually seen as a design variable under the control of the investigator. However, in some studies it is uncontrolled, and in others the actual plot size differs from that asserted. These observations lead us to some conclusions and recommendations on the use of occupancy as a state variable.

A SPATIAL MODEL OF OCCUPANCY

We start with a simple model that makes explicit the relationship between animal locations, animal activity, plot size, and occupancy of plots. We assume that animals exist within home ranges with a definite boundary and a stable location in space. The set of home-range centers may be viewed as a realization of a stochastic spatial point process (Fig. 1). The focus of this

paper is plot-based asymptotic occupancy, which we will denote ψ_p. A plot is occupied if it is overlapped by any home range, although at any instant there may be no individual in a particular occupied plot (following Tyre et al. 2003). Given that a plot is occupied, the cumulative probability of at least one individual actually being present approaches 1.0 asymptotically, over time. Methods for incomplete detection generally rely on replicate independent samples from multiple times or subplots; the asymptotic definition is appropriate because the probability of unavailability due to temporary absence is absorbed in the probability of detection (assuming independence between samples).

We focus on circular plots and home ranges with constant radius that are distributed at random, because in this case there is a simple expression for the expected occupancy:

$$E(\psi_p) = 1 - \exp[-\pi(r_{HR} + r_p)^2 D], \qquad (1)$$

where r_p and r_{HR} are the respective radii and D is the population density (i.e., the intensity parameter of the Poisson distribution of home-range centers). This derives from the zero term of the Poisson distribution for the number of individuals in a circle of radius $r_p + r_{HR}$, and may also be seen as a special case of the 'hitting' rate for a Boolean model with circular grains (Stoyan and Mecke 2005). The equivalent expression for a rectangular plot of length l and breadth b ($l = b$ for a square plot) is

$$E(\psi_p) = 1 - \exp[-(lb + 2(l + b)r_{HR} + \pi r_{HR}^2)D]. \qquad (2)$$

Results derived from the idealized model do not necessarily apply when plots are non-circular or home ranges are under- or over-dispersed relative to a Poisson distribution; we used simulation to assess the robustness of inferences to such deviations.

PLOTS OF KNOWN AREA

Equation 1 provides insight on the relative effects of plot size and home-range size. For a given absolute population density D, $E(\psi_p)$ is determined by the sum of the radii, and the value is equally sensitive to variation in either radius on an absolute scale. However, the intuition of
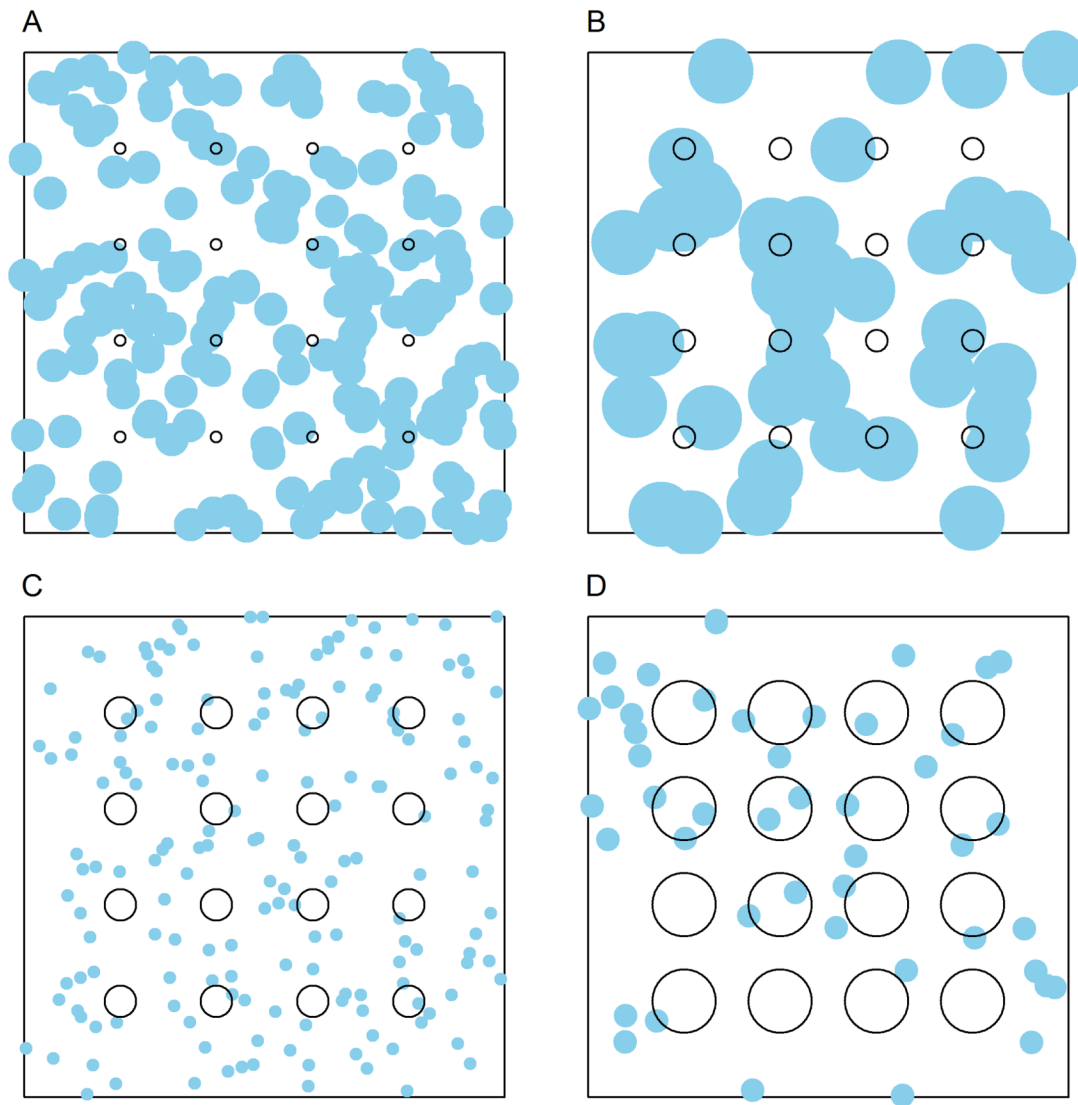
Fig. 1. Realizations of four scenarios with identical expected plot occupancy (≈ 0.684) and varying plot radius r_p, home range radius r_{HR}, and density D in an area 30 units square. The proportion of area covered by home ranges is 0.467 in (A) and (B) and 0.076 in (C) and (D). (A) $r_p = 1/(2\sqrt{2})$, $r_{HR} = 1$, $D = 0.20$; (B) $r_p = 1/\sqrt{2}$, $r_{HR} = 2$, $D = 0.05$; (C) $r_p = 1$, $r_{HR} = 1/(2\sqrt{2})$, $D = 0.20$; (D) $r_p = 2$, $r_{HR} = 1/\sqrt{2}$, $D = 0.05$.

ecologists cited in the Introduction relates to the *relative* sizes of plots and home ranges, and Eq. 1 may be rearranged to give

$$E(\psi_p) = 1 - \exp[-n(R + \sqrt{R} + 1)], \qquad (3)$$

where R is the ratio of home-range area to plot area, and $n = \pi r_p^2 D$ is the expected number of home-range centers within a plot. For square plots $\sqrt{R}$ is replaced by $4\sqrt{(R/\pi)}$. We graph these relationships for various levels of n in Fig. 2. Plot

shape (circular vs. square) has a trivial effect. When home ranges are small relative to plot size ($R < 0.01$), ψ_p is determined largely by plot radius and population density; ψ_p is strongly influenced by home-range size for at least some $R > 0.01$, although the effect depends on density and is truncated by saturation at 100% occupancy.

Deviations from random dispersion of home ranges change the predicted occupancy for a
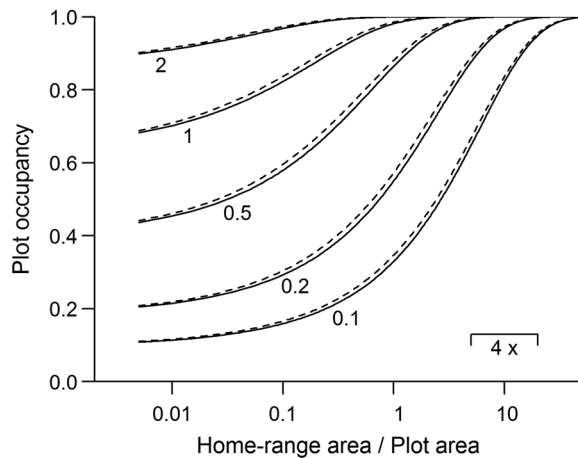
Fig. 2. Expected plot occupancy as function of relative home-range size for various densities, expressed as expected number per plot (n in range 0.01–2.0). Randomly dispersed circular home ranges; continuous curves are for circular plots (Eq. 1) and dashed curves for square plots (Eq. 2). The bottom line, for example, represents the expected plot occupancy over a range of relative home-range sizes when a plot is expected to contain 0.01 animals at any instant. Scale bar indicates a 4-fold variation in home range size with plot size held constant.

given density and relative plot size. We simulated data to determine the effect of maximal dispersion (non-overlapping, exclusive home ranges) and one level of clumping (see the Appendix for details and the Supplement for computer code). The patterns of Fig. 2 were largely maintained when home ranges were exclusive: differences in predicted occupancy were large (>10%) only for high density or large home ranges (Fig. 3). There is greater scope for deviation from the ideal pattern when home ranges are aggregated, and occupancy tends to a very small value in the extreme case that all home ranges coincide in space. For constant and moderate aggregation the pattern nevertheless parallels that for random dispersion (Fig. 3).

POINT SAMPLES WITH PASSIVE DETECTORS

As plot size approaches zero, occupancy comes to be determined by the product of density and home-range area, and by the dispersion of home ranges. The relationship for randomly dispersed

home ranges follows from Eq. 1, and results for other cases may be obtained by simulation. Across the range of cases considered (home ranges exclusive, random or mildly aggregated) the effect of dispersion is not large (Fig. 4).

POINT SAMPLES WITH DETECTION AT A DISTANCE

Surveys conducted by a stationary observer from a point generally differ from point samples with short-range passive detectors because animals are detected at a distance, by sight or sound. Although such surveys are point-based, the area within which animals may be detected is a finite plot. If detections are included only when they lie within a fixed radius, and detection probability is positive for all distances within this radius, then the point survey is equivalent to searching a circular plot of known size. When these conditions do not hold the plot size is uncertain. Uncertainty with regard to plot size has two related effects: it means the survey does not measure a well-defined state variable, and measurements of 'occupancy' are conditional on factors that influence detection distance (species, vegetation density, weather, background noise, observer acuity, etc.). This defeats the purpose of repeat sampling to estimate detection probability, as comparisons of occupancy over time or space may still be confounded with any variable that influences detection.

In practice, the condition that detection probability should be positive everywhere within the fixed radius is inadequate. Significant decline in detection probability with distance within the nominal radius of the plot induces strong heterogeneity in individual detection probability (Efford and Dawson 2009). Heterogeneity, in turn, causes bias in estimates of occupancy that may be so extreme that the measured occupancy bears little relation to the nominal plot size, as may be shown by simulations (Fig. 5). The magnitude of the bias may be represented by 'effective plot size', the size of plot that would yield estimates like those observed, under an idealized scenario with Poisson-distributed home ranges and no radial decline in detection probability. We can use Eq. 1 to determine this as a fraction of the nominal plot size (A):
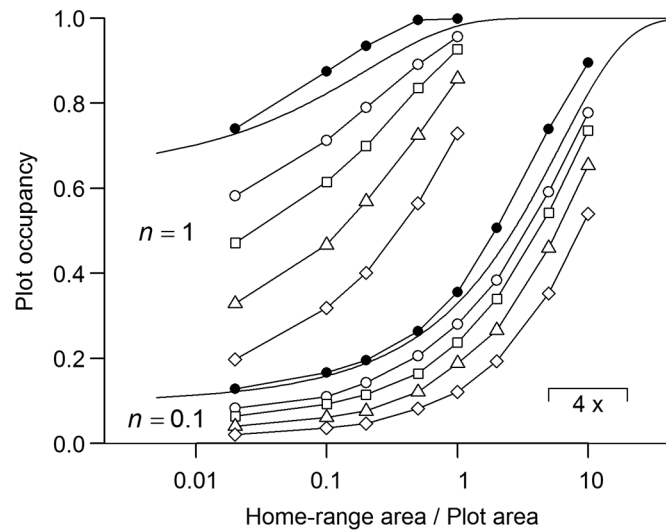
Fig. 3. Effect of home-range dispersion on expected plot occupancy as function of relative home-range size. Simulated results (methods in Appendix A) for circular home ranges at two levels of density expressed as expected number per plot ($n = 0.1$ and $n = 1.0$). Filled circles, exclusive home ranges; open symbols, clustered home ranges with mean cluster sizes 1 (circle), 2 (square), 4 (triangle) or 8 (diamond) (cluster scale σ = home-range radius; see Fig. 4 for example). Continuous curves show expected plot occupancy for randomly dispersed ranges as in Fig. 2. Scale bar indicates a 4-fold variation in home range size.

$$A_e = \frac{\log(1 - \mathrm{E}[\hat{\psi}])}{\log(1 - \psi)} A.$$

We used simulation to determine the effective plot size for sampling in a circular plot of nominal radius w when the per-sample probability of detecting an individual declined radially according to a half-normal function with scale σ (Efford and Dawson 2009). With the standard occupancy estimator, A_e ranged from $0.19A$ to $0.22A$ for $\sigma = 0.2w$ ($0.4 \leq \psi_p \leq 0.9$), and from $0.67A$ to $0.72A$ for $\sigma = 0.4w$ ($0.2 \leq \psi_p \leq 0.9$), the smaller A_e being associated with larger ψ_p. The results are affected slightly by the estimator used for occupancy; for example the Royle and Nichols (2003) estimator with a Poisson mixture resulted in A_e from $0.21A$ to $0.23A$ for $\sigma = 0.2w$, and from $0.73A$ to $0.75A$ for $\sigma = 0.4w$ over the same ranges of ψ_p as before. Variation in A_e across the simulated range of true occupancy was relatively minor, given a constant detection function and estimator.

FINITE PLOTS WITH POINT SAMPLES FOR SPATIAL REPLICATION

Sampling of replicated subplots is one way to obtain samples for fitting models that adjust for imperfect detection. Attention has focused recently on two aspects: modeling serial dependence of samples along a transect (Hines et al. 2010, Guillera-Arroita et al. 2011), and the effect of sampling with or without replacement (Kendall and White 2009, Guillera-Arroita 2011). We used simulation to evaluate the performance of two standard estimators with spatially replicated samples when the population follows our basic conceptual model.

Animals were assumed to occupy overlapping and randomly dispersed circular home ranges. Density was varied to achieve expected plot-level occupancy of $\psi_p = 0.1$, 0.3, 0.5, 0.7 or 0.9, using Eq. 2. Home-range radius R was 0.2, 0.4 or 0.8 times the grid-cell side s (area ratio $\approx$ 0.13, 0.50, 2.01). Sampling was simulated at random points ($N = 5$, 10) within each of 100 grid cells on a single occasion. Both the standard estimate and
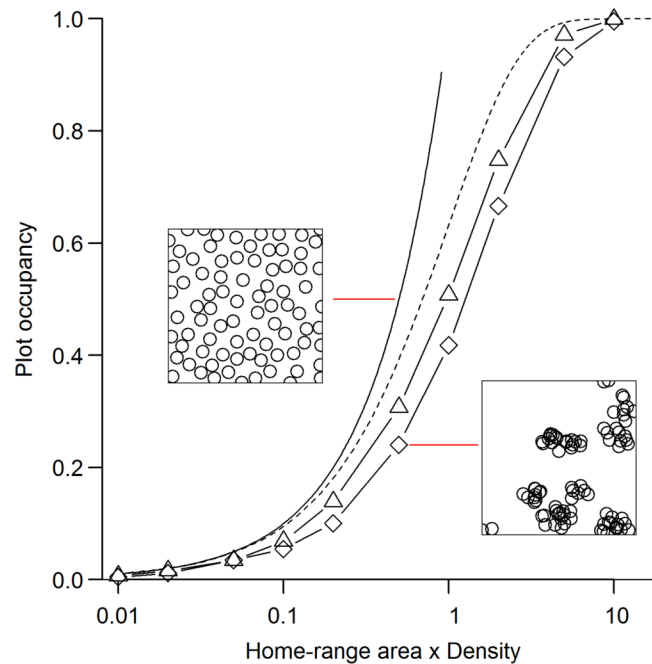
Fig. 4. Effect of home-range size, density and dispersion on occupancy at passive point detectors, given circular home ranges. Analytical results for exclusive home ranges (solid curve) and random (Poisson-distributed) home ranges (dashed curve). Simulated results (see methods in Appendix) with mean cluster sizes of 4 (triangle) and 8 (diamond), and bivariate-normal within-cluster dispersion (σ = radius of home range). The closest packing of circular home ranges is achieved when (area $\times$ density) = $\pi\sqrt{3}/6 \approx 0.91$.

the Poisson-based occupancy estimate of Royle and Nichols (2003) were computed for each of 1000 replicates at each combination of input levels. The performance of estimators was assessed in terms of bias ($\hat{\psi} - \psi$). The bias of the naïve estimator (the observed proportion of points at which the species was present) was used as a reference.

Modeling of incomplete detection was most effective at reducing bias when occupancy was high and home ranges were small (Fig. 6). The Royle-Nichols estimator was uniformly less biased than the standard estimator, but remained strongly biased (>20% relative bias) in most scenarios. Increasing the number of points sampled from 5 to 10 reduced bias noticeably only in the two scenarios with very low numbers of detections ($R = 0.2s$, $\psi_p = 0.1, 0.3$). Even the Royle-Nichols estimator performed poorly in most scenarios with $R > 0.2s$, in the sense that bias was reduced by less than 50% of the bias in the naïve estimates.

OCCUPANCY AS A SURROGATE FOR POPULATION DENSITY

Using one state variable as a surrogate for another presumes a consistent relationship between them that is independent of other factors likely to vary during monitoring. We graphed contours of ψ_p as a function of the combined radius ($r_{HR} + r_p$) and density for the model with random dispersion (Fig. 7A). The relationships in Fig. 7 are easier to grasp for specific linear units: if these are kilometers then the x-axis ranges from 0–300 m and the y-axis from 0–1 ha^{-1}. For fixed radius (a vertical section through Fig. 7A), increasing density causes an increase in ψ_p. For fixed density (a horizontal section through Fig. 7A), any increase in radius also increases ψ_p. The two contributing variables are confounded: constant ψ_p is compatible with widely varying combinations of combined radius and density (along a contour), whereas varying occupancy may or may not imply varying density. Note also the region above and to the right of the 90%
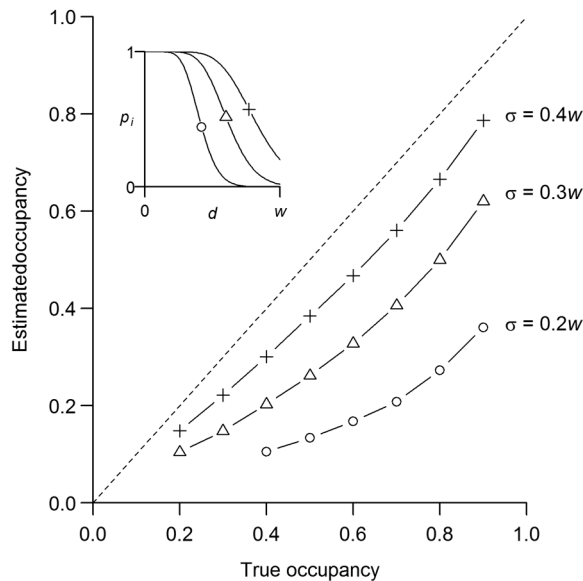
Fig. 5. Effect of radial decline in the detection probability of individuals on standard occupancy estimates from circular plots of radius w. Sampling was simulated at 100 independent points over 5 occasions. Per-occasion detection was a half-normal function of distance from the survey point with scale parameter σ; inset shows the individual detection probability, aggregated over 5 occasions, as a function of distance. Bias declines as σ/w increases, approaching zero (dashed diagonal) for $\sigma/w > 1$.

contour in which ψ_p saturates and responds very little to variation in either density or radius.

Variation in the dispersion of home ranges also affects ψ_p at a given density. Exclusive circular home ranges may be packed without overlap up to a maximum density of $\sqrt{3}/(6r_{HR}^2)$, and over this range the proportion of area covered by home ranges is $\pi r_{HR}^2 D$, with maximum value $\pi\sqrt{3}/6 \approx 0.907$. The resulting contours of area covered (Fig. 7B) are closer than in the random case, and the zone of saturation is enlarged.

DISCUSSION

We investigated the significance of plot size and home-range size for the design of studies using occupancy as a state variable, focusing on asymptotic occupancy of plots in continuous habitat. Asymptotic occupancy is the parameter generally measured by 'snapshot' sampling with either temporal or spatial replication when plots are not closed to cross-boundary home-range movements. When plots are very large or very small, either plot size or home-range size dominates, and we start by discussing these cases.

Very large plots

We consider plots to be 'very large' when they exceed 100 times the area of a home range, or 10 times home-range area when density is low (Fig. 2). Expected occupancy is only weakly affected by home-range size when plots are very large, as shown by the flatness of curves at the left of Fig. 2. Potential issues of confounding are therefore minor as long as plot size is clearly defined and constant. We emphasize that estimates of plot occupancy are specific to the plot size; as this comprises the unit of measurement it should be reported with each estimate. The term 'proportion of area occupied' (PAO) invites misinterpretation and should be avoided for plot-based estimates (note that the proportion of area covered by home ranges is 0.467 in Fig. 1A and B, and 0.076 in Fig. 1C and D, while plot occupancy remains constant at 0.684). Over- and under-dispersion relative to a Poisson distribution have predictable quantitative effects on occupancy, but as long as dispersion remains constant the qualitative patterns persist (Fig. 3). Although there are few conceptual problems with large plots, there may be practical difficulties if the habitat cannot accommodate many plots, or plots are too large to be searched with equal intensity. It is probably for these reasons, and because it is costly to search large areas, that very large plots are rarely used in occupancy studies.

Very small plots

At the other extreme are plots that are essentially point samples, as may be obtained with passive devices such as traps or cameras. Fig. 2 is not useful for these as the density expressed as n per plot tends to zero. Expected occupancy depends on the product of home-range area and density, with overlaid effects of dispersion as before (Fig. 4). Point detectors sample the combined 'footprint' of the population formed from the union of the home ranges (Fig. 1). Asymptotic occupancy is then the
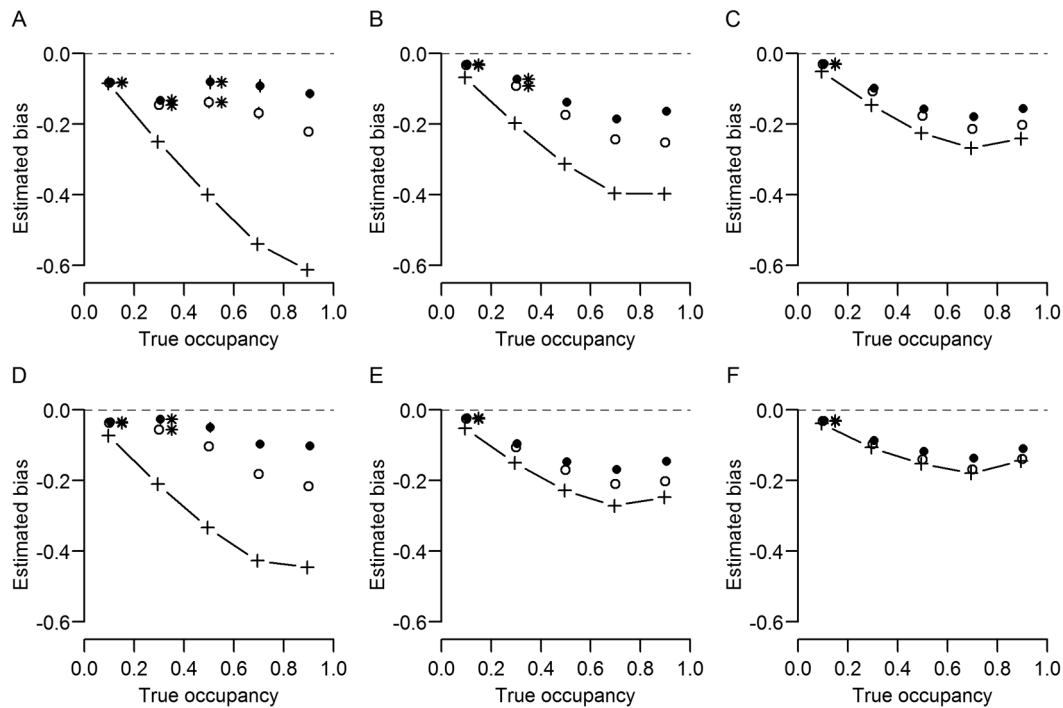
Fig. 6. Bias of occupancy estimates using repeated point sampling within grid cells. Crosses, naïve occupancy; open circles, standard estimate; filled circles, Royle-Nichols Poisson model estimate. Asterisks mark estimates that have been adjusted by dropping data sets in which the species was detected at only one point within any grid cell (these invariably result in an occupancy estimate of ~1.0). Bars (obscured by symbols) indicate 95% confidence intervals for the mean of 1000 simulations.
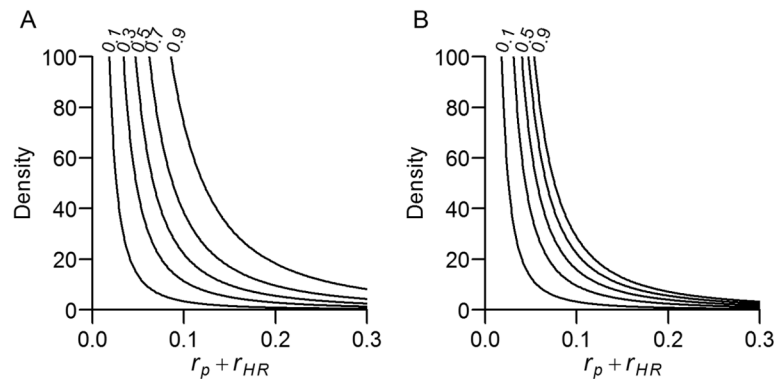

Fig. 7. Expected fractional coverage of the plane by disks of varying radius ($r = r_p + r_{HR}$) and density D. (A) randomly located centers ($p = 1 - \exp(-\pi r^2 D)$) (B) exclusive ranges ($p = \pi r^2 D$ given $D \leq \sqrt{3}/(6r^2)$). Points along a contour share the same expected coverage. Radius and density are assumed to be measured in compatible units (e.g., km and km^{-2}). Rightmost contour in (B) is close to the maximum coverage of hexagonally packed disks (≈ 0.907).

proportion of habitat that lies within at least one home range. We suggest that use of the term PAO should be restricted to this measure when habitat is continuous. PAO is plot-free (indepen-

dent of any arbitrarily chosen plot size); however, users should be conscious that it is a joint function of density and home-range size (PAO tends to zero as home ranges shrink).

MacKenzie and Royle (2005) preferred the term 'proportion of area used' to 'proportion of area occupied' when home-range movements dominated over plot size. This terminology has been adopted only sporadically in the literature, and has lately been confused by the application of 'use' to mean availability for detection (Mordecai et al. 2011).

The definition of PAO becomes unclear if home ranges do not have a definite edge. Statistical models of animal activity such as the bivariate normal imply that home ranges are unbounded, and field data support the general 'fuzziness' of boundaries (e.g., Powell 2000). A possible response to this objection is to assert that within a finite time interval only part of the home range will be visited, and occupancy may be defined with an arbitrary temporal constraint (e.g., 'probability plot visited within 24 hours'). This is true, but unhelpful, as current estimation models are stochastic and, under the assumption of random temporary emigration, the quantity estimated is the asymptotic occupancy. The critical assumption that home ranges are bounded deserves more attention from practitioners, especially those applying occupancy methods to data from traps or cameras.

Plots of intermediate size

Many sampling designs use plots that are similar in size to home ranges, within 1–2 orders of magnitude. In some cases this is deliberate. For example, Wintle et al. (2005) defined plots for acoustic surveys of owls and arboreal marsupials on the basis of the distance at which species may be detected, which they note was similar to the home-range radius. In other cases, investigators have chosen plots to be larger than a home range, but the plots were not large enough to escape the effects of home-range size on occupancy, as judged by our results (e.g., Karanth et al. 2011). Plots in this intermediate size range are more prone to problems of interpretation than very large or very small plots. Poor definition of plot size is a particular problem, as we discuss in the next section.

Plots of indeterminate size

The boundaries of plots that have been surveyed for occupancy are often ill-defined. This applies particularly to strip transects of unspecified width (e.g., Field et al. 2005), and to point-based samples of unspecified radius (for example, a baited trap may lure animals from an unknown distance, or an observer may record all birds heard at a point without determining their distances). In practice this means that plot size is determined by the observation conditions, and that occupancy is no longer a well-defined state variable.

Protocols that discard detections beyond a certain radius do not satisfactorily define plot size if the radius is too large. In multi-species studies, there may be many species for which the probability of detecting an individual falls to near zero within the nominal plot radius. A prominent example is the North American Breeding Bird Survey (BBS; Robbins et al. 1986) in which the nominal radius (400 m) is several times the maximum detection distance of most species (Norvell et al. 2003, Alldredge et al. 2007).

For a particular set of unbounded or nominally bounded point samples there will be an effective plot size A_e, as we have documented for selected radial detection functions by simulation. However, A_e is an unknown that cannot be determined from non-spatial occupancy data (we determined it by simulation with known parameter values), and expressing bias in this way tends to obscure the fact that plot size remains indeterminate.

The issue of variable effective plot size appears to have been overlooked and deserves more critical attention in the design of field studies. New statistical methods for occupancy are routinely exemplified with analyses of BBS data (Royle and Nichols 2003, MacKenzie et al. 2004, Dorazio et al. 2006, Royle 2006, Royle and Link 2006, Kendall and White 2009) or other point observations for which the effective plot radius is unknown (e.g., Dupuis et al. 2010), which may lead practitioners to think this is an appropriate use.

"Multi-scale" occupancy

The robust sampling design of Pollock (1982) may be adapted to estimate components of

occupancy (Nichols et al. 2008, Mordecai et al. 2011). We discuss the skunk (*Mephitis mephitis*) example of Nichols et al. (2008) and O'Connell and Bailey (2011). They randomly located one station with multiple detectors within each cell in a random subset of cells from a 100-m × 100-m grid. The model yielded estimates of both the probability of presence at a station at a particular time conditional on overall presence (θ), and the conventional parameter for overall presence at each station (ψ). Nichols et al. (2008) attributed θ and ψ to different spatial scales, and interpreted ψ as the probability that a *grid cell* was occupied. In that case we would expect ψ to increase with increasing size of grid cell. However, as all data pertaining to a grid cell are collected from a single station under their sampling design, the estimate of ψ is unrelated to grid-cell size. We therefore suggest that the components θ and ψ should not be interpreted in terms of occupancy at two spatial scales (cf. Nichols et al. 2008, O'Connell and Bailey 2011). It is better to think of the parameters as relating to different temporal scales: $\psi\theta$ is the probability that a station is occupied at a particular time (instantaneous occupancy), whereas ψ represents the overall (asymptotic) occupancy of stations. This interpretation is also appropriate if availability for detection changes in situ (e.g., when an animal switches between vocalizing and silent) rather than because animals move relative to the station.

In this paper we have emphasized the scale-dependence of plot occupancy, using theoretical models such as Eq. 1 to predict variation with plot size. Empirical verification in the face of imperfect detection would seem to require replicate surveys at each spatial scale, rather than just one.

Persistent sign

Recording ephemeral sign (tracks, feces, etc. that do not persist between samples) is functionally equivalent to direct observation of animals, but surveys of persistent sign (long-lived tracks, feces, antler rubs, etc.) raise other issues. If it can be assumed that signs remain fixed throughout the survey, then repeated searches for sign on plots of any size allow estimation of the 'perceptual' component of incomplete detection and hence the instantaneous proportion of plots containing sign. This number will depend not only on plot size, population density and home-range size, as with asymptotic occupancy, but also on the processes by which signs are generated and decay. Given that a plot is within a home range, the probability that it contains sign (whether detected or undetected) is a form of 'availability' analogous to the instantaneous probability an animal is present in a plot overlapping its home range. It is tempting to assume that sign is pervasive, i.e., present with certainty in any plot overlapping a home range, but such cases are probably rare. If sign is not pervasive, occupancy is confounded with spatial or temporal variation in the processes controlling the availability of signs (e.g., rates of defecation and fecal decay).

Occupancy as a state variable in population monitoring

Occupancy has been considered an alternative to direct measures of population size or density when the resources for monitoring are limited (MacKenzie and Nichols 2004). Confounding of ψ_p with home-range size and plot size creates the potential for serious inferential error or loss of inferential power when ψ_p is used as a surrogate for density in population monitoring. The risk of relying on occupancy when plots are small relative to home-range size was clearly flagged by MacKenzie and Royle (2005:1108): "If occupancy is being employed as a surrogate for abundance estimation, the level of 'use' [PAO] may be irrelevant and even misleading." This has not deterred many users.

Inferential error results when effects of interest, such as population trend over time, are confounded with extraneous variables such as home-range size or detection distance. The theoretical potential for confounding is shown in Fig. 7, and there are biological reasons to expect these variables to co-vary with population density and with habitat type. Negative covariation of home-range size and population density is widespread (for taxonomically diverse examples see Wilson 1975, Abramsky and Tracy 1980, Schoener and Schoener 1982, and Trewhella et al. 1988); under these conditions density may increase without a detectable change in occupancy. Differences between habitats in the dispersion and availability of resources may drive variation

in the size or overlap of home ranges independently of density (e.g., Adams 2001); in this case ψ_p may suggest a difference in density when there is none. Species interactions also may cause extraneous variation in apparent occupancy through their effects on spatial behavior (e.g., reduced movement in the presence of predators) or detection (e.g., acoustic saturation). Vegetation density can strongly affect the maximum distance at which individuals can be detected (e.g., Bibby and Buckland 1987); large effects, which may be common, potentially cause ψ_p to increase while density declines, or vice versa. Species often differ markedly in home-range size and detection distances; differences in ψ_p between species are therefore unlikely to be interpretable when plots are small.

Heterogeneity

The probability of detecting each individual presumably increases with the area of overlap between its home range and a plot, and the number of individuals also varies among occupied plots. As a result, the probability of detecting the species given presence is expected to be heterogeneous among plots. Non-spatial occupancy models produce biased estimates, as we showed for spatially replicated point sampling when home ranges were of similar size to the plot. A non-spatial Poisson mixing model for p (Royle and Nichols 2003, Royle 2006) allows for heterogeneity in detection probability due to varying number, and slightly reduces the bias. The persistent component of bias in this case appears to be due to heterogeneity among plots in the proportion of area occupied.

Potential for enhanced models

Can the limitations of occupancy as a state variable be corrected? The method may be applied safely to designs that use repeated searches of large, well-defined plots. In other cases it becomes necessary to model home-range movements and, in many cases, detection at a distance. Given data on marked individuals, the methods of spatially explicit capture–recapture (SECR) (e.g., Efford et al. 2009) may be used to estimate a detection function that combines these effects, and leads directly to an estimate of density (Borchers and Efford 2008). The cost of data collection for SECR likely is considerably

higher than for occupancy surveys, so this is barely a solution to the underlying problem. However, Conroy et al. (2008) suggested combining widespread, low-cost presence-absence sampling with intensive capture–recapture sampling at selected locations. Their approach was non-spatial and used independent parameterizations of detection for the presence-absence and capture–recapture phases. A spatially explicit model of two-phase sampling could have significant advantages over the approach of Conroy et al. (2008) in that it would yield an estimate of population density with a minimum of parameters while accounting for individual heterogeneity due to location. While this approach would use presence-absence data, the state variable (density) would not suffer from the problems associated with occupancy. We note also that the use of incomplete presence-absence data to estimate and map species distributions (e.g., Sargeant et al. 2005) does not suffer from the problems we have highlighted when occupancy is used as a state variable.

Conclusions

We suggest that the confounding effects of plot size and home-range characteristics (size, dispersion) on occupancy have received insufficient attention from practitioners, and that the range of situations in which ψ_p is a robust parameter for population monitoring is smaller than previously believed. The effect is especially damaging when plot size is unknown, a situation that is surprisingly common in published applications of occupancy, particularly in relation to bird point counts, but also when animals are lured to cameras or traps (e.g., Thorn et al. 2009). The growing popularity of occupancy in monitoring programs (Marsh and Trenham 2008, Morrison et al. 2008) is due in part to its statistical endorsement by respected researchers (MacKenzie and Nichols 2004, MacKenzie et al. 2005, Nichols et al. 2009) and the apparent ease with which it can be estimated from historical data. However, the perception that plot-based occupancy estimates from non-spatial models are resistant to confounded variation in detection probability may be mistaken. The estimates provide a reliable index of population change only when the effective plot size, home-range size, and home-range dispersion can be assumed constant, and

home ranges are bounded. Whether the assumptions are justified in a particular study cannot be determined from internal evidence, as occupancy surveys do not generally provide information on home ranges or detection distances. In this respect, occupancy has much in common with the surrogates for abundance commonly called indices (e.g., Anderson 2003, Johnson 2008). The arguments against ψ_p as a surrogate for population density are those against indices in general: it is implausible that the relationship with density is constant in the long term, and in any particular study the relationship is usually unknown. Like other indices, occupancy should be used judiciously, taking care to control for possible confounding during data collection and analysis (Johnson 2008).

ACKNOWLEDGMENTS

LITERATURE CITED

Abramsky, Z., and C. R. Tracy. 1980. Relation between home range size and regulation of population size in *Microtus ochrogaster*. Oikos 34:347–355.

Adams, E. S. 2001. Approaches to the study of territory size and shape. Annual Review of Ecology and Systematics 32:277–303.

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

Anderson, D. R. 2003. Response to Engeman: index values rarely constitute reliable information. Wildlife Society Bulletin 31:288–291.

Baddeley, A., and R. Turner. 2005. Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software 12:1–42.

Bailey, L. L., J. E. Hines, J. D. Nichols, and D. I. MacKenzie. 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecological Applications 17:281–290.

Bibby, C. J., and S. T. Buckland. 1987. Bias of bird census results due to detectability varying with habitat. Acta Oecologica 8:103–112.

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

Conroy, M. J., J. P. Runge, R. J. Barker, M. R. Schofield, and C. J. Fonnesbeck. 2008. Efficient estimation of abundance for patchily distributed populations via two-phase, adaptive sampling. Ecology 89:3362–3370.

Dorazio, R. M., J. A. Royle, B. Söderström, and A. Glimskär. 2006. Estimating species richness and accumulation by modeling species occurrence and detectability. Ecology 87:842–854.

Dupuis, J. A., F. Bled, and J. Joachim. 2010. Estimating the occupancy rate of spatially rare or hard to detect species: a conditional approach. Biometrics 67:290–298.

Efford, M. G., D. L. Borchers, and A. E. Byrom. 2009. Density estimation by spatially explicit capture–recapture: likelihood-based methods. Pages 255–269 *in* D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Springer, New York, New York, USA.

Efford, M. G., and D. K. Dawson. 2009. Effect of distance-related heterogeneity on population size estimated from point counts. Auk 126:100–111.

Field, S. A., A. J. Tyre, K. H. Thorn, P. O'Connor, and H. P. Possingham. 2005. Improving the efficiency of monitoring by estimating detectability: a case study of foxes, *Vulpes vulpes*, on the Eyre Peninsula, South Australia. Wildlife Research 32:1–6.

Guillera-Arroita, G. 2011. Impact of sampling with replacement in occupancy studies with spatial replication. Methods in Ecology and Evolution 2:401–406.

Guillera-Arroita, G., M. S. Ridout, and B. J. T. Morgan. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1:131–139.

Guillera-Arroita, G., B. J. T. Morgan, M. S. Ridout, and M. Linkie. 2011. Species occupancy modeling for detection data collected along a transect. Journal of Agricultural, Biological, and Environmental Statistics 16:301–317.

Hines, J. E., J. D. Nichols, J. A. Royle, D. I. MacKenzie, A. M. Gopalaswamy, N. S. Kumar, and K. U. Karanth. 2010. Tigers on trails: occupancy modeling for cluster sampling. Ecological Applications 20:1456–1466.

Johnson, D. H. 2008. In defense of indices: the case of bird surveys. Journal of Wildlife Management 72:857–868.

Karanth, K. U., A. M. Gopalaswamy, N. S. Kumar, S. Vaidyanathan, J. D. Nichols, and D. I. MacKenzie. 2011. Monitoring carnivore populations at the landscape scale: occupancy modelling of tigers from sign surveys. Journal of Applied Ecology 48:1048–1056.

Kendall, W. L. 1999. Robustness of closed capture–recapture methods to violations of the closure assumption. Ecology 80:2517–2525.

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.

MacKenzie, D. I., and J. D. Nichols. 2004. Occupancy as a surrogate for abundance estimation. Animal Biodiversity and Conservation 27:461–467.

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. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Academic Press, San Diego, California, USA.

MacKenzie, D. I., J. D. Nichols, N. Sutton, K. Kawanishi, and L. L. Bailey. 2005. Improving inferences in population studies of rare species that are detected imperfectly. Ecology 86:1101–1113.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

MacKenzie, D. I., J. A. Royle, J. A. Brown, and J. D. Nichols. 2004. Occupancy estimation and modeling for rare and elusive populations. Pages 149–172 *in* W. L. Thompson, editor. Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, D.C., USA.

Marsh, D. M., and P. C. Trenham. 2008. Current trends in animal and plant population monitoring. Conservation Biology 22:647–655.

Mattsson, B. J., and M. R. Marshall. 2009. Occupancy modeling as a framework for designing avian monitoring programs: a case study along Appalachian streams in southern West Virginia. Pages 617–632 *in* T. D. Rich, C. Arizmendi, D. W. Demarest, and C. Thompson, editors. Proceedings of the Fourth International Partners in Flight Conference, McAllen, Texas, February 13–16, 2008. http://www.partnersinflight.org/pubs/McAllenProc/articles/PIF09_Monitoring/Mattsson_PIF09.pdf.

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.

Morrison, M. L., W. M. Block, M. D. Strickland, B. A. Collier, and M. J. Peterson. 2008. Wildlife study design. Second edition. Springer, New York, New York, USA.

Nichols, J. D., L. L. Bailey, A. F. O'Connell, N. W. Talancy, E. H. C. Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. Journal of Applied Ecology 45:1321–1329.

Nichols, J. D., L. Thomas, and P. B. Conn. 2009. Inferences about landbird abundance from count data: recent advances and future directions. Pages 201–235 *in* D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Springer, New York, New York, USA.

Norvell, R. E., F. P. Howe, and J. R. Parrish. 2003. A seven-year comparison of relative-abundance and distance-sampling methods. Auk 120:1013–1028.

O'Brien, T. G., J. E. M. Baillie, L. Krueger, and M. Cuke. 2010. The Wildlife Picture Index: monitoring top trophic levels. Animal Conservation 13:335–343.

O'Connell, A. F., and L. L. Bailey. 2011. Inference for occupancy and occupancy dynamics. Pages 191–205 *in* A. F. O.'Connell, J. D. Nichols, and K. U. Karanth, editors. Camera traps in animal ecology. Springer, Tokyo, Japan.

Pollock, K. H. 1982. A capture–recapture design robust to unequal probability of capture. Journal of Wildlife Management 46:752–757.

Powell, R. A. 2000. Animal home ranges and territories and home range estimators. Pages 65–110 *in* L. Boitani and T. K. Fuller, editors. Research techniques in animal ecology. Columbia University Press, New York, New York, USA.

R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robbins, C. S., D. Bystrak, and P. H. Geissler. 1986. The Breeding Bird Survey: its first fifteen years, 1965–1979. U.S. Fish and Wildlife Service, Resource Publication 157. Washington, D.C., USA.

Royle, J. A. 2006. Site occupancy models with heterogeneous detection probabilities. Biometrics 62:97–102.

Royle, J. A., and W. A. Link. 2006. Generalized site occupancy models allowing for false positive and false negative errors. Ecology 87:835–841.

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

Sargeant, G. A., M. A. Sovada, C. C. Slivinski, and D. H. Johnson. 2005. Markov chain Monte Carlo estimation of species distributions: a case study of the swift fox in western Kansas. Journal of Wildlife Management 69:483–497.

Schoener, T. W., and A. Schoener. 1982. Intraspecific variation in home-range size in some Anolis lizards. Ecology 63:809–823.

Stoyan, D., and J. Mecke. 2005. The Boolean model: from Matheron till today. Pages 151–182 *in* M. Bilodeau, F. Meyer, and M. Schmitt, editors. Space, structure and randomness. Contributions in honor of Georges Matheron in the fields of geostatistics,

random sets, and mathematical morphology. Springer, New York, New York, USA.

Thorn, M., D. M. Scott, M. Green, P. W. Bateman, and E. Z. Cameron. 2009. Estimating brown hyaena occupancy using baited camera traps. South African Journal of Wildlife Research 39:1–10.

Trewhella, W. J., S. Harris, and F. E. McAllister. 1988. Dispersal distance, home-range size and population density in the red fox (*Vulpes vulpes*): a quantitative analysis. Journal of Applied Ecology 25:423–434.

Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecological Applications 13:1790–1801.

Weir, L., I. J. Fiske, and J. A. Royle. 2009. Trends in anuran occupancy from northeastern states of the North American Amphibian Monitoring Program. Herpetological Conservation and Biology 4:389–402.

Wilson, E. O. 1975. Sociobiology: the new synthesis. Harvard University Press, Cambridge, Massachusetts, USA.

Wintle, B. A., R. P. Kavanagh, M. A. McCarthy, and M. A. Burgman. 2005. Estimating and dealing with detectability in occupancy surveys for forest owls and arboreal marsupials. Journal of Wildlife Management 69:905–917.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16:446–453.

Zylstra, E. R., R. J. Steidl, and D. E. Swann. 2010. Evaluating survey methods for monitoring a rare vertebrate, the Sonoran desert tortoise. Journal of Wildlife Management 75:1311–1318.

SUPPLEMENTAL MATERIAL

APPENDIX

Simulations to determine the effect of variations in home-range dispersion on occupancy

Simulations used the R package 'spatstat' of Baddeley and Turner (2005) to generate realizations of exclusive and aggregated home-range patterns. Home ranges were simulated as circular disks (radius R) whose centers lay within an arbitrary window of area A large enough to include all disks potentially overlapping a single central plot. Disk centers were placed according to either a simple spatial inhibition process (function 'rSSI') with inhibition radius R or a Thomas spatial point process (function 'rThomas'). A Thomas process is a Neyman-Scott cluster process with circular bivariate normal within-cluster dispersion. The number of disks n was a Poisson random variable with $E(n) = D.A$ where D was the population density. The range of values simulated with the spatial inhibition algorithm was limited to avoid problems when not all n disk centers could be fitted within the window.

Simulated occupancy was the proportion of 10000 plots that were overlapped by at least one home range. In order to simulate a range of ratios of home-range area (πr_{HR}^2) to plot area (πr_p^2 for a circular plot radius r_p), we held the relevant plot dimension constant at 1.0 and varied r_{HR}.

An R (R Development Core Team 2011) package for these simulations and others is provided in the Supplement.

SUPPLEMENT

R package for conducting the simulations described in the main text (*Ecological Archives* C003-004-S1).