

Species-specific differences in detection and occupancy probabilities help drive ability to detect trends in occupancy

ROBIN STEENWEG,^{1,4,†} MARK HEBBLEWHITE,¹ JESSE WHITTINGTON,² AND KEVIN MCKELVEY³

¹Wildlife Biology Program, W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, Montana 59812 USA

²Parks Canada, Banff National Park Resource Conservation, Banff, Alberta T1L 1K2 Canada

³US Forest Service, Rocky Mountain Research Station, Missoula, Montana 59801 USA

Citation: Steenweg, R., M. Hebblewhite, J. Whittington, and K. McKelvey. 2019. Species-specific differences in detection and occupancy probabilities help drive ability to detect trends in occupancy. *Ecosphere* 10(4):e02639. 10.1002/ecs2.2639

Abstract. Occupancy-based surveys are increasingly used to monitor wildlife populations because they can be more cost-effective than abundance surveys and because they may track multiple species, simultaneously. The design of these multi-species occupancy surveys affects statistical power to detect trends in occupancy because individual species vary in resource selection, detection probability, and rarity. We tested for differences in the ability of a large-scale monitoring program to detect changes in single-species occupancy of 13 medium-large mammal species captured on $n = 183$ cameras systematically placed across five national parks in the Canadian Rockies ($\sim 21,000$ km²). We focus the interpretation of our findings on three species at risk: grizzly bear, wolverine, and caribou. We found that statistical power to monitor trends in occupancy depends not only on the established elements associated with power (sampling size, effect size, and variation in estimates), but also on species-specific detection and occupancy probabilities. These two probabilities, however, affected power differently. For most species in our study, power is insensitive to detection probability. Increasing replicate-specific detection probability only improved power when the cumulative detection probability was below 0.80. Therefore, efficient species monitoring must consider that power no longer improves by increasing sample size or the replicate-specific detection probability once this threshold is reached. On the other hand, species with occupancy probabilities close to 0.5 had lower statistical power than those with higher or lower occupancy, that is, power was higher for both rare and very common species. This pattern is due to the heretofore-underappreciated effect of the binomial variation in occupancy. The implications of these findings are species-specific. Grizzly bears, for example, had high detection and occupancy probabilities, resulting in high power to detect a population change. Conversely, wolverines had low detection probability and the power to detect change could be improved if detection probability was increased using lure or complimentary survey techniques. Caribou, however, with both low detection and occupancy probabilities, were likely too rare on the landscape to rely on camera-based occupancy for monitoring. Practitioners should be aware of these species-specific trade-offs and may need to tailor monitoring programs to prioritize particular species of conservation concern.

Key words: caribou (*Rangifer tarandus*); carnivores; grizzly bear (*Ursus arctos*); multi-species monitoring; occupancy; remote cameras; Rocky Mountains; statistical power; trend detection; wolverine (*Gulo gulo*).

Received 1 January 2019; accepted 11 January 2019; final version received 8 February 2019. Corresponding Editor: Debra P. C. Peters.

Copyright: © 2019 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

⁴ Present address: Alberta Environment and Parks, Grande Prairie, Alberta T8V 6J4 Canada.

† **E-mail:** robin.steenweg@gov.ab.ca

INTRODUCTION

With limited conservation funding, there is a need for affordable multi-species monitoring (Simberloff 1998). Tracking spatial distribution as a state variable using occupancy models (MacKenzie et al. 2002) is more cost-effective than tracking abundance or population dynamics and has been advocated by many practitioners (O'Brien et al. 2010, Noon et al. 2012, Steenweg et al. 2017). For many species, occupancy sampling design usually focuses on one or a few focal species (Burton et al. 2015). Many occupancy data collection methods, for example, using remote cameras, acoustic surveys, eDNA, or snow tracking, are multi-species by nature, detecting many non-target species even when directed at a focal species. Capitalizing on these non-target data presents an opportunity to maximize the effectiveness and efficiency of multi-species monitoring methods. Not only do non-target detections represent a wealth of information collected at little-to-no additional cost, but these data can additionally be used to improve detection probability modeling for rare (i.e., low occupancy) species (Iknayan et al. 2014). As with any monitoring program, design trade-offs must be considered (Yoccoz et al. 2001) and statistical power analysis provides a method to compare how design choices affect the ability to detect population trends over time (Steidl et al. 1997). With multi-species monitoring, there will be differential trade-offs when monitoring trends of target versus non-target species.

Effective sampling designs for monitoring occupancy need to consider the ecology of the individual species (Burton et al. 2015). Territoriality and spatial overlap between individuals, for example, can negatively bias occupancy estimates if the occupancy model relies on the assumption that individuals are not restricted to discrete habitat (Efford and Dawson 2012). The effect of spatial sampling scale on occupancy estimates, and thus power to detect a change in the population occupancy, depends on the relationship between spatial sampling grain and home-range size (MacKenzie and Nichols 2004, Noon et al. 2012). When using aerial plot sampling to track wolverine occupancy, statistical power increases with smaller spatial grain (Ellis et al. 2014). The effect of sampling grain on

occupancy estimates, however, depends on the sample type (e.g., plots, transects, or points), where estimates of occupancy are scale-invariant when using point samples (Steenweg et al. 2018). Temporal sampling extent (i.e., duration of sampling) can also drive the definition of occupancy when sampling mobile organisms; definitions can vary from ephemeral use to annual occupancy, depending on survey length (Steenweg et al. 2018). With rigorous sampling design, the power of an occupancy-monitoring program should be robust to different spatial patterns of occupancy decline (e.g., decline from periphery, due to an ecological trap, or decline from range edge; Steenweg et al. 2016a). Furthermore, most occupancy studies are designed for one target or focal species, raising uncertainty about their suitability to monitor trends of non-target species (Nichols 2010). While some previous studies have published closed-form and simulation-based methods to calculate statistical power for occupancy data, and have provided advice for study-design trade-offs between allocation of effort among spatial and temporal replicates (e.g., MacKenzie and Royle 2005, Bailey et al. 2007, Guillera-Arroita and Lahoz-Monfort 2012; see details below), there are few explicit empirical tests of trade-offs when monitoring multiple species, especially in one system. Advice for sampling design and analysis may differ among species for which data are collected simultaneously.

Statistical error can be summarized into two error types. Type I error, α , is the probability of falsely detecting an increase or decrease in population occupancy when it has, in truth, not changed. Type II error, β , is the probability of missing a population change that is occurring. Power, $1-\beta$, is the probability of correctly rejecting the hypothesis that the population is not changing, thus supporting the hypothesis that there is a population trend. There is a trade-off between these two errors, and in the context of monitoring threatened species, the consequences of failing to detect a real decline may have consequences that are more difficult to recover from (e.g., increased extinction risk with long time lags for recovery) than those associated with a false alarm (e.g., short-term financial cost; Field et al. 2005). Monitoring programs must balance these trade-offs when designing studies to collect data.

Effective sampling designs for monitoring occupancy also rely on statistical design elements relevant to all monitoring projects. The three main components of calculating statistical power include the desired effect size that a program could detect, sampling size, and variance in parameter estimates (Steidl et al. 1997). The first component, desired effect size, is determined by management objectives, which need to balance logistical and financial costs against the risk of missing an important population change. When using occupancy models that correct for imperfect detection, the second component, sample size (i.e., number of locations), must also be balanced with the number of repeated visits (MacKenzie and Royle 2005, Bailey et al. 2007, Guillera-Aroita et al. 2010, Guillera-Aroita and Lahoz-Monfort 2012). In occupancy modeling, the general advice for allocating effort among sites and replicate visits to those sites is to increase the number of replicates for common species and increase the number of sites for rare species (MacKenzie and Royle 2005). Specifically, if one considers the cumulative detection probability (p^*), which is the probability of detecting a species at least once during the entire survey, a monitoring program should plan to repeat sampling until $p^* > 0.8$ (Long et al. 2008) or near 0.9 (McKann et al. 2013), above which there are only modest increases in statistical power. Similarly, some studies on statistical power in occupancy trend suggest that at least 60 sites are required to estimate a trend in occupancy (McKann et al. 2013, Steenweg et al. 2016a) and 120 sites to estimate colonization and extinction rates (McKann et al. 2013). Another design consideration is which statistical test will be used to test for changes in occupancy. Sampling the same sites across years and testing for changes in occupancy using a paired t -test will have higher power than sampling different sites and testing for changes in occupancy using a Wald test (Steenweg et al. 2016a). Finally, variance of parameter estimates will be affected by sample size, sampling error, and natural variability over space and time of the parameter of interest, here, occupancy. A key step in power analysis to guide monitoring programs is to estimate variability, which is species-specific, and then determine how to

minimize this variability and thus maximize power.

Few studies, if any, have investigated the consequences of these relationships on statistical power to detect trends in occupancy in a multi-species setting. In this paper, we investigate statistical power to detect trends in occupancy of multiple medium and large mammals where sampling design was tailored to grizzly bear (*Ursus arctos*). Our research question is as follows: In a multi-species community where monitoring setup for a target/focal species (grizzly bears), how do species-specific detection (p) and occupancy (ψ) probabilities affect abilities to detect changes in occupancy across different species? To answer this question, first we estimated simple occupancy estimates of large mammals in our community who span a range of p and ψ probabilities. Second, with these estimates, we calculated statistical power to detect a decline for each species using the closed-form equations from Guillera-Aroita and Lahoz-Monfort (2012). Third, we hypothesize that, when sampling conditions are held constant among species, both species-specific p and ψ will affect power systematically, simply due to mathematical properties of the occupancy estimator. We predict that these systematic effects will transcend ecological differences among species. To test this hypothesis, we analyzed data on 13 species of large mammals collected simultaneously from $n = 183$ remote cameras in a Rocky Mountains setting where a previous study showed high power to detect trends in grizzly bear occupancy (Steenweg et al. 2016a). This suite of mammals includes species that are common and species that are rare (i.e., with known low occupancy, abundance, or both). As well, this list includes species that are commonly seen in the open and that are cryptic. Finally, we generalize these findings to all combinations of p and ψ probabilities for individual species.

MATERIALS AND METHODS

Study area

The Canadian Rockies study area (~21,000 km²) includes five national parks, Jasper, Banff, Yoho, Kootenay, and Waterton Lakes national parks, (JNP, BNP, YNP, KNP, and WLNP; Steenweg et al. 2016a). The topography

and climate are extreme, ranging from 1000 to 3800 m and -50°C to 35°C , respectively. The full compliment of large mammals that were present pre-contact continues to roam these protected areas, including grizzly bears, black bears (*U. americanus*), wolves (*Canis lupus*), coyotes (*C. latrans*), red fox (*Vulpes vulpes*), cougars (*Puma concolor*), lynx (*Lynx canadensis*), bobcat (*L. rufus*), wolverine (*Gulo gulo*), elk (*Cervus canadensis*), white-tailed deer (*Odocoileus virginianus*), mule deer (*O. hemionus*), moose (*Alces americanus*), bighorn sheep (*Ovis canadensis*), and mountain goats (*Oreamnos americanus*). The only exceptions include bison (*Bison bison*), which were recently (in 2017) reintroduced into Banff (Steenweg et al. 2016b), and caribou (*Rangifer tarandus*), which are still present in Jasper, but have been recently extirpated from Banff (Hebblewhite et al. 2009). See Steenweg et al. (2016a) for more study area details.

We collected occupancy data for 13 species of medium–large mammal (>10 kg) using remote cameras. In 2012, we deployed $n = 183$ cameras, at a density of one camera per 100 km^2 in JNP, BNP, YNP, and KNP. Our sampling design was focused on monitoring grizzly bears, with greater than one camera per average grizzly bear home range, which are $\sim 520\text{ km}^2$ and $\sim 1405\text{ km}^2$ for females and males, respectively (Stevens and Gibeau 2005). This systematic sampling nearly covered each National park, and in the smaller WLNP, we were able to increase the density of cameras to 1 per 25 km^2 . Within cells, we tried to maximize detection probability through placing cameras on human or wildlife trails (often at junctions) and near bear rub trees whenever possible (Burton et al. 2015). Camera models included covert motion-triggered cameras (Hyperfire and Rapidfire models; Reconyx, Holmen, Wisconsin, USA) and a few visible glow cameras (Silent Image, Reconyx; Panthera V4, Panthera, Bronx, New York, USA). We attached cameras to trees at waist height, pointed slightly downwards, and set them to take five pictures per movement trigger with no delay between triggers. For public safety consideration in national parks, we did not use bait or lure. We serviced cameras three to four times a year, and we classified all camera images using program Timelapse (Greenberg and Godin 2015).

Occupancy analysis

Occupancy models that correct for imperfect detection require repeated sampling across time, space, or using multiple observers (MacKenzie et al. 2018). Cameras were deployed continuously, and we, therefore, discretized the data for occupancy analysis, but the choice of sampling-replicate length often varies across studies (Burton et al. 2015). Previously, we showed that with camera data, shorter time intervals improve precision on occupancy estimates (Steenweg et al. 2016a), and thus, we divided the data into 26 one-week intervals (total time, May 1–October 31). We also show results for analyzing the same data as 6 one-month intervals. Given the long sampling duration of our study, the occupancy parameter for most species should be interpreted as the probability of occupancy at any point during the summer by any number of individuals. For some species with large home ranges (larger than our 100 km^2 spatial sampling scale), the occupancy parameter is better interpreted as the probability of *use* at any point during the summer by any number of individuals (Steenweg et al. 2018). We did not include any covariates on either p or ψ in the occupancy models (using, thus, $p(\cdot)$ $\psi(\cdot)$ models). This strategy may have produced biased estimates due to unmodeled heterogeneity but maximized our objective to compare power among species without difficulties in interpretation due to different covariates among species. All analyses were carried out in the unmarked package (Fiske and Chandler 2015) in the statistical software R (R Development Core Team 2016).

Power analysis

Power to detect changes in occupancy can be estimated using stand-alone programs (Bailey et al. 2007), simulation (Steenweg et al. 2016a), or closed-form algebraic equation (Guillera-Arroita and Lahoz-Monfort 2012). To reduce complexities and permit easy comparisons across species, we used closed-form equations where the asymptotic variance, σ^2 , of an occupancy estimate, ψ , can be calculated as (MacKenzie and Royle 2005, Guillera-Arroita and Lahoz-Monfort 2012).

$$\sigma^2 = \frac{\psi}{S} \left\{ (1 - \psi) + \frac{1 - p^*}{p^* - kp(1 - p)^{k-1}} \right\} \quad (1)$$

S is the number of sites sampled, k is the number of sampling replicates, and p is the detection probability, which is sampling-replicate-specific. The cumulative detection probability, p^* , is the probability of detecting a species at least once during k sampling replicates and is calculated as follows:

$$p^* = 1 - (1 - p)^k \quad (2)$$

which assumes that all detection events are independent. The variance Eq. 1 has two components: the variance of a binomial proportion, $\psi(1 - \psi)/S$, and a correction for imperfect detection. Modifying Eq. 4 from Guillera-Aroita and Lahoz-Monfort (2012), we calculated power, G , for each species as

$$G = 1 - \beta = \left\{ 1 - \Phi \left(\frac{z_\alpha \sqrt{\sigma_1^2 + \sigma_2^2} - (\psi_1 - \psi_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right) \right\} \quad (3)$$

where $z_\alpha \sqrt{\sigma_1^2 + \sigma_2^2}$ is the boundary of the critical region for a one-tailed Wald test, and Φ is the cumulative distribution function for the normal distribution. We maintained Type I error at $\alpha = 0.05$ and used the conventional 80% power as our threshold to qualify adequate power (Gelman and Hill 2006). We then simulated absolute declines in occupancy between two sampling years of up to 0.30 and calculated when power had exceeded 80%, following Steenweg et al. (2016a). We then determined the minimum decline that we were able to detect with 80% power. We calculated the minimum decline we could detect for the 13 mammals detected within our study area, given their estimated occupancy and detection probabilities. We then generalized these power calculations across all possible parameter combinations but restricted to 0.05–0.95 (at intervals of 0.05) for detection probability and 0–0.9 (at intervals of 0.1) for occupancy. Parameter ranges were restricted due to issues of estimation close to boundaries (0 or 1; MacKenzie et al. 2018; see *Discussion*). To explore how the choice of sampling-replicate length affects p and potentially p^* , we

repeated these analyses but discretized the same data into 6 one-month sampling-replicate intervals instead of 26 one-week replicates. Finally, we also altered the spatial and temporal sampling replicates (n and k , respectively) in simulations in Appendix S3.

RESULTS

Occupancy (ψ) and detection probabilities (p) varied across species in our study area (Fig. 1, Table 1). Both ψ and p were highest for common species and lowest for rare species. Occupancy ranged from 0.79 (95% CI: 0.72–0.85) for grizzly bear to 0.04 (0.02–0.08) for caribou. Detection probability was highest for black bears (0.25; 0.23–0.27) and lowest for wolverine (0.06; 0.04–0.09). These species-specific differences in ψ and p had varying effects on statistical power (Fig. 2, Table 1). We plotted the power curves for four species (black bear, wolverine, grizzly bear, and red fox) to assess how our range of estimates for p and ψ affected power. Of these four species, we had the highest power to detect trends of red fox, with an 80% probability to detect a decline of

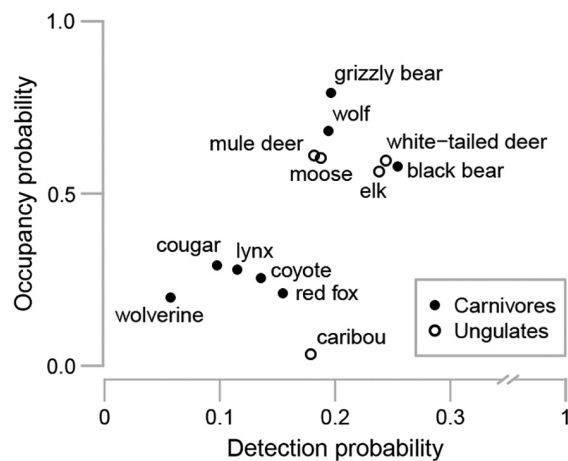


Fig. 1. Point estimates of summer occupancy and detection probabilities for 13 mammals across the Canadian Rocky Mountain study area (see Table 1 for confidence intervals). Estimates were based on data from $n = 183$ cameras for 26 weeks (May 1–October 31) and using occupancy models with no covariates. Note that detection probabilities depended on our chosen replicate interval length of seven days; see Table 1; Appendix S2 for more details.

Table 1. Effects of species-specific occupancy and detection probabilities (mean estimates and 95% confidence intervals) on the minimum decline in occupancy that the current sampling design could detect with 80% power and $n = 183$ cameras.

Discretization	Min. decline detected		Occupancy probability (ψ)		Detection probability (p)		Cumulative p (p^*)	
	6 months	26 weeks	6 months	26 weeks	6 months	26 weeks	6 months	26 weeks
Grizzly bear	0.131	0.114	0.82 (0.74–0.88)	0.79 (0.72–0.85)	0.48 (0.44–0.52)	0.20 (0.18–0.21)	0.980	0.997
Wolf	0.145	0.126	0.68 (0.60–0.75)	0.68 (0.61–0.75)	0.48 (0.44–0.53)	0.19 (0.18–0.21)	0.981	0.996
Mule deer	0.152	0.129	0.65 (0.56–0.73)	0.61 (0.53–0.69)	0.41 (0.36–0.46)	0.18 (0.17–0.20)	0.956	0.995
Moose	0.15	0.129	0.63 (0.55–0.71)	0.60 (0.53–0.68)	0.44 (0.39–0.49)	0.19 (0.17–0.21)	0.968	0.996
WT deer	0.148	0.129	0.62 (0.54–0.70)	0.60 (0.52–0.67)	0.47 (0.43–0.52)	0.24 (0.23–0.26)	0.978	0.999
Black bear	0.147	0.129	0.60 (0.52–0.67)	0.58 (0.50–0.65)	0.51 (0.46–0.56)	0.25 (0.23–0.27)	0.986	1
Elk	0.148	0.129	0.59 (0.51–0.67)	0.56 (0.49–0.64)	0.48 (0.43–0.52)	0.24 (0.22–0.26)	0.979	0.999
Cougar	0.139	0.116	0.30 (0.22–0.40)	0.29 (0.22–0.38)	0.29 (0.23–0.37)	0.10 (0.08–0.12)	0.874	0.931
Lynx	0.13	0.111	0.29 (0.22–0.37)	0.28 (0.21–0.36)	0.34 (0.27–0.41)	0.12 (0.09–0.14)	0.914	0.959
Coyote	0.123	0.105	0.27 (0.20–0.35)	0.26 (0.19–0.33)	0.36 (0.29–0.44)	0.14 (0.11–0.16)	0.934	0.977
Red fox	0.12	0.096	0.24 (0.17–0.32)	0.21 (0.15–0.28)	0.33 (0.25–0.41)	0.15 (0.13–0.18)	0.908	0.987
Wolverine	0.159	0.113	0.22 (0.13–0.37)	0.20 (0.12–0.31)	0.17 (0.10–0.28)	0.06 (0.04–0.09)	0.672	0.786
Caribou†	NA	NA	0.03 (0.01–0.07)	0.04 (0.02–0.08)	0.42 (0.23–0.64)	0.18 (0.12–0.26)	0.962	0.994

Notes: The same summer 2012 camera data (May 1–October 31) were discretized into 26 one-week sessions and 6 one-month sessions. Data from this table were used to create Fig. 1.

† Occupancy rates for caribou were too low to reliably compute the power to detect change, see *Discussion* for details.

0.096. Meanwhile, we could sufficiently detect declines in black bear occupancy of 0.129 (Fig. 2). We were unable to calculate the level of decline we could detect for caribou due to estimation problems near boundaries for binomial data (see *Discussion*). For all species, when discretizing the same data into longer sampling-replicate windows (1-month, rather than 1-week), detection probability estimates increased (Table 1; Appendix S2: Fig. S1); however, occupancy estimates were largely insensitive to this change in sampling-replicate window; values were slightly higher with 1-month windows and precision decreased (Table 1; see also Steenweg et al. 2016a,b). Due to the decreased precision of the estimates, power to detect change also decreased across all species with 1-month windows (Table 1; Appendix S2: Fig. S2).

To understand these effects of ψ and p on power more clearly and expand the parameter space beyond the 13 species we tested, we plotted the minimum absolute decline in occupancy that can be detected with adequate power for all combinations of ψ and p (Fig. 3; see *Materials and Methods*). Except when detection probability is extremely low ($p < 0.1$), power to detect change is insensitive to changes in detection probability, because the sampling period (26 weeks) was

sufficiently long to ensure a high cumulative detection probability, p^* (see *Discussion* for more details). When detection probability approaches 0, for all levels of occupancy, power is severely degraded. Power is lowest for species that exhibit both extremely low detection probabilities and high occupancies (Fig. 3; Appendix S1: Fig. S1). The power to detect change is, however, sensitive to changes in occupancy with the lowest power associated with occupancy rates of 0.5 and power increasing as species become either very common or very rare; the highest power to detect declines occurs in species with low occupancy and detection probabilities >0.1 (Fig. 3; Appendix S1: Fig. S1). Of the species we sampled and were able to compute power, only the wolverine had a replicate-specific detection probability low enough (0.06, 1-week detection intervals; Table 1) to fall into that region in which low detection probabilities significantly decreased power. See also Appendix S1 for additional perspective of the 3D plot in Fig. 3, and see Appendix S1 for R code to adjust parameters and recreate Fig. 3 with the ability to rotate to any perspective. The shape of the 3D plot was robust to changes in spatial and temporal sampling replicates (n and k , respectively; see Appendix S3).

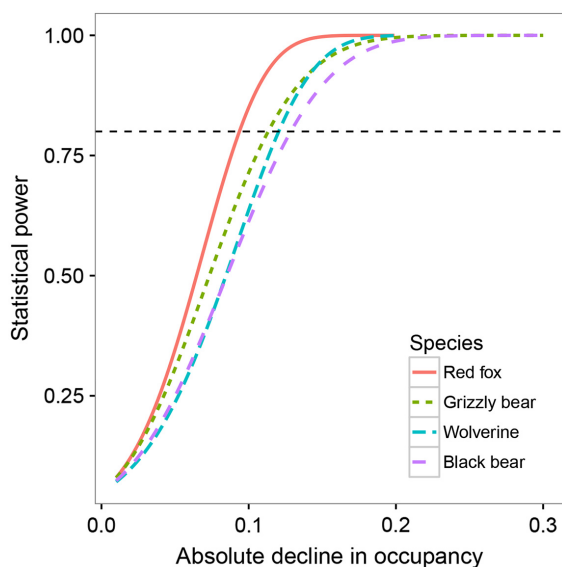


Fig. 2. Statistical power for detecting an absolute decline in occupancy. Data collected from $n = 183$ cameras deployed for 26 weeks (May 1–October 31) in the Canadian Rocky Mountains. Power for each species was calculated using a closed-form equation. The horizontal dashed line represents a conventional threshold of 80% power.

DISCUSSION

We found that the ability of our large-scale multi-species monitoring program to detect changes in population occupancy of 13 large mammal species depended not only on the usual elements associated with statistical power (sampling size, effect size, and variation in estimates), but also on both species-specific detection (p) and occupancy (ψ) probabilities. Furthermore, power to detect a change was affected differently by these two parameters (Fig. 3), each resulting in different study-design considerations. Except for wolverines, p was high enough for all species such that power was insensitive to differences in p , because sample size was large and p^* was high. We had the highest power to detect trends for relatively rarer species, such as lynx and cougar, despite their lower p because, although scarce, p was adequate to ensure p^* , the cumulative probability of detecting a species at least once during the entire survey was above 0.95 in most cases (Table 1). This result highlights the importance of ensuring

$p^* > 0.8$, remarked upon by others (Long et al. 2008). Because of the length of time that the cameras were recording, the broad differences between p of grizzly bear (1-week, $p = 0.79$) compared to red fox (1-week, $p = 0.21$) only resulted in a change of 0.01 in p^* given our sampling time. For wolverines, on the other hand, this was not the case: p was low enough to reduce power. As a result, despite a similar occupancy estimate to red fox, the 1-week p^* was 0.2 lower for wolverine, resulting in lower power (Table 1). Visually, the effects of reduced p on power occurred at $p < 0.11$ (Fig. 3). Wolverine was the only species in our study to intersect this back upslope (saddle) of the p - ψ parameter space beyond this inflection point (Fig. 3). Using longer replicate-sampling length (1-month rather than 1-week intervals) did increase p (Table 1; Appendix S1: Fig S2), but not power, because there was no corresponding increase in p^* (Table 1; Appendix S1: Fig S3). We emphasize here that given Eq. 2, maximizing p^* can be achieved through either increasing p (e.g., using lure, or placing cameras on trails such as in this study) or increasing the number of replicates, k . For camera-based monitoring projects, increasing k entails leaving cameras out longer, which affects the interpretation of the estimated occupancy parameter (Steenweg et al. 2018), but there are other ways to increase detection probability and power as well.

Even for species for which occupancy estimates based on one sampling design provide insufficient statistical power, additional steps can be taken to improve statistical power through combining different data collection methods to increase detection. Wolverines are a Species of Concern in Canada (COSEWIC 2014) and listed as Data Deficient in Alberta (AB SRD 2003). We had low statistical power to detect trends in wolverine in our study area because of their low detection probability. With our study design and sample size ($n = 183$ cameras), however, we could still detect a change in occupancy of 0.159, despite no use of covariates in our model to potentially improve precision, and thus power. Using lure or bait is one method to increase p , and correspondingly, p^* , for species caught on cameras (Burton et al. 2015) although their use may affect inferences (MacKenzie et al. 2018). Detection probability for wolverine, for example, can be greatly increased with beaver carcasses

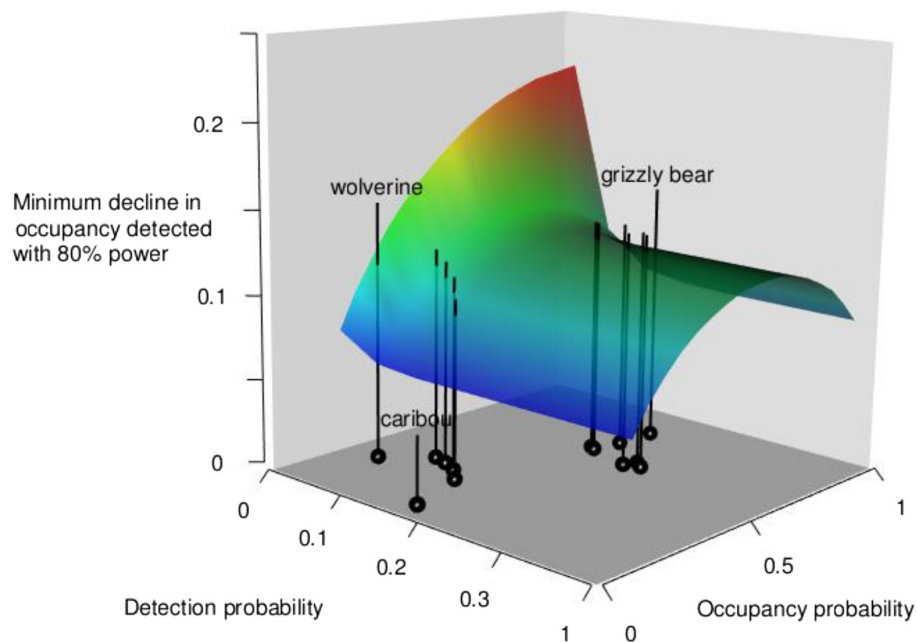


Fig. 3. Minimum decline in occupancy that can be detected with 80% power between two years, as a function of both detection and occupancy probabilities. Cooler colors indicate higher statistical power. Declines in occupancy were simulated using $n = 183$ cameras deployed for 26 weeks. The bottom plane is equivalent to Figure 1, showing estimates from 13 species collected May 1–October 31 in the Canadian Rocky Mountain study area. See Appendix S1: Fig. S1 for more perspectives of this 3D plot.

(Fisher et al. 2013, Heim 2015). When logistical constraints prohibit frequent reapplication of lure or bait, other data sources and methods could also be used to improve statistical power, such as snow tracking. Whittington et al. (2015) demonstrated, for example, that despite low p , snow tracking could be used to monitor trends of wolverine and lynx occupancy with high statistical power, further emphasizing that monitoring programs need to look beyond p .

When the focus is shifted away from estimating parameters to estimating statistical power, we see that maximizing p^* does not necessarily maximize power; the binomial nature of ψ must also be considered. Once the $p^* \approx 0.8$ threshold is reached, the resulting statistical power function is dominated by the effect of ψ . The functional form of the power calculations (Fig 3.) shows that across the p - ψ parameter space, the concave shape of the relationship with power remains the same across values of p where $p^* > 0.8$. Power is consistently lower for species with ψ closer to 0.5. Rare species intersect the plane where the

concave shape drops as ψ approaches 0 and thus have higher power; the same is true for common species when ψ approaches 1 (Fig. 3). This concave shape is due to the heretofore-underappreciated nature of the variance of a binomial proportion, $\psi(1-\psi)/S$ (see *Materials and Methods*), where the variance is greatest at $\psi = 0.5$. For most of the parameter space, the effect of imperfect detection in Eq. 1 approaches 0. For example, despite black bears having the highest detection probability, our ability to detect a decline in black bear distribution was lower than many other species (Table 1). Other studies have shown that grizzly bears are more common than black bears in this area (Serrouya 1999, Sawaya et al. 2012). In our study, black bears had an occupancy estimate of 0.58, whereas grizzly bears had an occupancy of 0.79 (Table 1). As a result, despite lower detection probability, the grizzly bear occupancy estimate was further from $\psi = 0.5$ than black bears, resulting in higher power to detect trends in grizzly bear occupancy than in black bear occupancy.

Furthermore, although power to detect trends in occupancy is higher for species with occupancy estimates that are further from $\psi = 0.5$, extreme estimates near 0 or 1 have estimation issues (low precision and likely some bias) near those boundaries (see also Appendix S3). Using Eq. 1 to calculate power for caribou, which had a very low occupancy estimate of 0.04 (CI: 0.02–0.08), resulted in being able to detect a decline of 0. This impossible result is due to a breakdown in Eq. 3 close to boundaries of 0 and 1, which is common for binomial data. Caribou were only detected at 6 of 183 camera sites. There is little consensus in the statistical literature about how to estimate precision of binomial proportions near boundaries (e.g., ψ near 0 or 1). A general statistical rule of thumb states that for binomial data, the number of successes and the number of failures must each be at least 5 (Brown et al. 2001), and others recommend at least 10 (Steve Kanters, *personal communication*). For this reason, we restricted our power calculations to parameter space of 0.1–0.9 for ψ , and 0.05–0.95 for p . For very rare species, and correspondingly very common species, monitoring programs may have difficulty estimating occupancy and detecting a decline due to the binomial nature of occupancy. Additional or alternative (e.g., targeted-) monitoring methods may be required for rarer species observed during multi-species programs.

Paradoxically, power may be higher for non-target species than for target species. In this example, our sampling designed to detect changes in grizzly bear distribution produced similar power to detect change for an additional 11 species, only not for caribou where occupancy levels were so low that change metrics were not calculable. We have shown that p^* can be influenced by increasing the number of temporal replicates or, in the case of continuously recording camera sets, by sufficiently increasing the amount of recording time to ensure p^* is consistently >0.8 . This sampling intensity allows robust change detection to be calculated for many non-target species. For the target species, the study areas may be drawn to optimize the likelihood of encompassing a particular population of interest, and habitats may be preferentially sampled where the target species is thought to occur. On the other hand, non-target

species may be scarce in the study area, but due to the increased power to detect trends at lower occupancy, the study may have higher power for these rare non-targets. Assuming a reasonable level of p^* , the relationships in Fig. 3 show that for a scarce non-target species, the power to detect change would exceed that of a more common target species. Of equal interest is that a lack of efficacy associated with efforts that are generally undertaken to increase detection probabilities for the target species could potentially compromise ability to detect trends. Fig. 3 suggests that this effort can be misplaced if detecting a change in population occupancy is the primary study objective. In other words, accepting low levels of occupancy can be better than trying to target species-specific habitat because that can increase the total number of sites occupied to moderate levels.

Although occupancy (ψ) and detection probabilities (p) varied widely across species, they appear positively correlated (Fig. 1). Most rare species, for example wolverine and lynx, have both low ψ and p (Fig. 1). On the other hand, common species, such as most ungulates, had both high ψ and p . The largest carnivores in our study system (wolves and bears) commonly used trails and therefore also had both high ψ and p . Both ψ and p are related to abundance (N) but in different ways. Occupancy has been described as “a simple, information-poor summary of abundance, or the poor man’s abundance” (Kery and Royle 2016). Essentially, $\psi_{\text{true}} = 0$ when $N = 0$, and $\psi_{\text{true}} = 1$ when $N > 0$. Furthermore, p increases when N increases (Royle and Nichols 2003). Therefore, the correlation between ψ and p is likely caused by this interrelationship with N . Specifically, the correlation seen here between ψ and p (Fig. 1) is likely due to landscape-level abundance, with which occupancy is positively related to (Kery and Royle 2016). In this study, the cameras recorded for 26 week, enough time for individuals to travel in from adjacent areas, and local individuals to leave. Royle and Nichols (2003) models the changes in detection associated with abundance within the cell, but here detection will be affected both by local abundance and by landscape-level abundance, for which occupancy serves as an index. More research is needed into the positive correlation of ψ and p , and how this might affect statistical power.

MANAGEMENT IMPLICATIONS

With multi-species monitoring technologies expanding, more programs are maturing to the stage where statistical power to detect trends should be evaluated. We have shown that programs using occupancy to monitor population trends should aim to have $p^* > 0.8$ and can maximize p^* and power by increasing p and/or k . Statistical power was not necessarily highest, however, for species with the highest p^* . Power also depends on ψ and was higher for many of the rarer species in our system. Other species that are more common than the species we sampled here could potentially have similarly high power as very rare species. Changes in composite occupancy indexes, such as the wildlife picture index (Beaudrot et al. 2016), therefore, will also depend on species rarity. A decline in WPI in one area, for example, could be driven by the decline of a few rare species or a few common species, or potentially the decline of more moderately common species. More recognition is needed by practitioners of the binomial nature of occupancy as a state variable, in contrast to Poisson-distributed abundance, and how the species-specific differences explored here affect statistical power to detect a trend in occupancy. Furthermore, more research is required to assess the statistical power of composite indexes of occupancy.

ACKNOWLEDGMENTS

We thank Paul Lukacs, Winsor Lowe, Mike Mitchell, Marco Musiani, and three anonymous reviewers for comments on a previous draft. We also thank those who helped collect camera data including Parks Canada resource conservation staff, Ellen Brandell, Leah Pengelly, and many volunteers. Funding was provided by Parks Canada, the University of Montana, Panthera Inc., Alberta Biodiversity Monitoring Institute, Alberta Parks, and the Yellowstone to Yukon Conservation Initiative.

LITERATURE CITED

- AB SRD. 2003. Alberta Sustainable Resource Development. Alberta's wolverine (*Gulo gulo*) : data deficient.
- 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.
- Beaudrot, L., J. Ahumada, T. O'Brien, P. Alvarez-Loayza, K. Boekee, A. Campos-Arceiz, D. Eichberg, S. Espinosa, E. Fegraus, and C. Fletcher. 2016. Standardized assessment of biodiversity trends in tropical forest protected areas: The end is not in sight. *PLoS Biology* 14:e1002357.
- Brown, L. D., T. T. Cai and A. DasGupta. 2001. Interval estimation for a binomial proportion. *Statistical Science* 26:101–117.
- Burton, A. C., E. Neilson, D. Moreira, A. Ladle, R. Steenweg, J. T. Fisher, E. Bayne, and S. Boutin. 2015. Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology* 52:675–685.
- COSEWIC. 2014. COSEWIC assessment and status report on the Wolverine *Gulo gulo* in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa. xi + 76 pp. Available: <https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/wolverine-status-2014.html>
- Efford, M. G. and D. K. Dawson. 2012. Occupancy in continuous habitat. *Ecosphere* 3:art32.
- Ellis, M., J. Ivan, and M. Schwartz. 2014. Spatially explicit power analyses for occupancy-based monitoring of wolverine in the U.S. Rocky Mountains. *Conservation Biology* 28:52–62.
- Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management* 69:473–482.
- Fisher, J. T., S. Bradbury, B. Anholt, L. Nolan, L. Roy, J. P. Volpe, and M. Wheatley. 2013. Wolverines (*Gulo gulo luscus*) on the Rocky Mountain slopes: natural heterogeneity and landscape alteration as predictors of distribution. *Canadian Journal of Zoology* 91:706–716.
- Fiske, I., and R. Chandler. 2015. Overview of unmarked: an R package for the analysis of data from unmarked animals. CRAN. <https://cran.r-project.org/web/packages/unmarked/index.html>
- Gelman, A., and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, New York, USA.
- Greenberg, S., and T. Godin. 2015. A tool supporting the extraction of angling effort data from remote camera images. *Fisheries* 40:276–287.
- Guillera-Aroita, G., and J. J. Lahoz-Monfort. 2012. Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods in Ecology and Evolution* 3: 860–869.

- Guillera-Aroita, G., M. S. Ridout, and B. J. Morgan. 2010. Design of occupancy studies with imperfect detection. *Methods in Ecology and Evolution* 1:131–139.
- Hebblewhite, M., C. White, and M. Musiani. 2009. Revisiting extinction in national parks: Mountain Caribou in Banff. *Conservation Biology* 24:341–344.
- Heim, N. A. 2015. Complex effects of human-impacted landscapes on the spatial patterns of Mammalian Carnivores. University of Victoria, Victoria, British Columbia, Canada.
- Iknayan, K. J., M. W. Tingley, B. J. Furnas, and S. R. Beissinger. 2014. Detecting diversity: emerging methods to estimate species diversity. *Trends in Ecology & Evolution* 29:97–106.
- Kery, M., and A. Royle. 2016. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 1: prelude and static models. Academic Press, London, UK.
- Long, R., W. J. Zielinski, R. Long, P. MacKay, W. Zielinski and J. Ray. 2008. Designing effective noninvasive carnivore surveys. Pages 8–44 in *Noninvasive survey methods for carnivores*. Island Press, Washington, D.C., USA.
- MacKenzie, D., and J. 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. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MacKenzie, D., J. Nichols, J. Royle, J. Pollock, L. Bailey, and J. E. Hines. 2018. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego, California, USA.
- MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology* 42:1105–1114.
- McKann, P. C., B. R. Gray, and W. E. Thogmartin. 2013. Small sample bias in dynamic occupancy models. *Journal of Wildlife Management* 77:172–180.
- Nichols, J. D. 2010. The wildlife picture index, monitoring and conservation. *Animal Conservation* 13:344–346.
- Noon, B. R., L. L. Bailey, T. D. Sisk, and K. S. McKelvey. 2012. Efficient species-level monitoring at the landscape scale. *Conservation Biology* 26:432–441.
- O'Brien, T., J. Baillie, L. Krueger, and M. Cuke. 2010. The Wildlife Picture Index: monitoring top trophic levels. *Animal Conservation* 13:335–343.
- R Development Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84:777–790.
- Sawaya, M. A., J. B. Stetz, A. P. Clevenger, M. L. Gibeau, and S. T. Kalinowski. 2012. Estimating grizzly and black bear population abundance and trend in Banff National Park using noninvasive genetic sampling. *PLoS ONE* 7:e34777.
- Serrouya, R. 1999. Permeability of the trans-Canada highway to black bear movements in the bow river valley of Banff National Park. University of British Columbia, Vancouver, British Columbia, Canada.
- Simberloff, D. 1998. Flagships, umbrellas, and keystones: Is single-species management passe in the landscape era? *Biological Conservation* 83:247–257.
- Steenweg, R., M. Hebblewhite, J. Whittington, K. McKelvey, and P. Lukacs. 2018. Sampling scales define occupancy and underlying occupancy-abundance relationships in animals. *Ecology* 99:172–183.
- Steenweg, R., J. Whittington, M. Hebblewhite, A. Forshner, B. Johnston, D. Petersen, B. Shepherd, and P. M. Lukacs. 2016a. Camera-based occupancy monitoring at large scales: power to detect trends in grizzly bears across the Canadian Rockies. *Biological Conservation* 201:192–200.
- Steenweg, R., M. Hebblewhite, D. Gummer, B. Low, and B. Hunt. 2016b. Assessing Potential Habitat and Carrying Capacity for Reintroduction of Plains Bison (*Bison bison bison*) in Banff National Park. *PLoS ONE* 11:e0150065.
- Steenweg, R., et al. 2017. Scaling up camera traps—monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment* 15:26–34.
- Steidl, R. J., J. P. Hayes and E. Schaubert. 1997. Statistical power analysis in wildlife research. *Journal of Wildlife Management* 61:270–279.
- Stevens, S., and M. L. Gibeau. 2005. Home range analysis. Pages 144–152 in S. Herrero, editor. *Biology demography, ecology and management of grizzly bears in and around Banff National Park and Kananaskis country: the final report of Eastern slopes Grizzly Bear Project*. Faculty of Environmental design, University of Calgary, Alberta, Calgary, Canada.

- Whittington, J., K. Heuer, B. Hunt, M. Hebblewhite and P. Lukacs. 2015. Estimating occupancy using spatially and temporally replicated snow surveys. *Animal Conservation* 18:92–101.
- 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.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2639/full>