20

Optimising broad-scale monitoring for trend detection: review and re-design of a long-term program in northern Australia

Luke D. Einoder, Darren M. Southwell, Graeme R. Gillespie, Alaric Fisher, José J. Lahoz-Monfort and Brendan A. Wintle

Summary

One of the biggest considerations when designing a threatened species monitoring program is deciding how to survey populations across space and time. Too little effort invested in surveys can mean that trends or important changes in populations may not be detected when they occur, and too much effort is a waste of resources that could otherwise be spent on management. Power analysis is a useful tool for assessing: (1) how likely it is that a monitoring program will detect important changes in a species distribution and/or abundance; (2) the level of sampling effort required to detect changes with a desired degree of confidence; and (3) the

whereabouts, timing and sampling method that achieves desired power at the lowest cost. This chapter reports on a power analysis that informed the evaluation and re-design of a long-term terrestrial vertebrate monitoring program in northern Australia. Modelling estimated the spatial variation in current occupancy for nine species and estimated how difficult they are to detect during surveys. Using this information, distribution changes and surveys for a suite of species with varying initial distribution and detectability were simulated. These simulations were used to explore the power of current and alternative monitoring designs at detecting changes in the nine species of interest. It was found that the current monitoring program has sufficient power (>80%) to detect large declines (>40% change in distribution over 15 years) for only the most common and easy-to-detect species. Finally, the survey design (i.e. the combination of survey frequency, survey duration, and number of sites) that maximises the statistical power to detect population trends was determined.

Introduction

A common goal of threatened species monitoring is to detect population changes over time. Achieving a high probability of detecting important changes requires decisions about the sampling method (e.g. live trapping, camera trapping, spotlighting), sampling duration (i.e. the number of days/nights spent surveying sites), sampling frequency (i.e. how regularly sites are surveyed), and the location and number of sites. Limited budgets usually impose trade-offs between these different components of a monitoring program. For example, sampling sites more regularly or increasing the time spent at sites may reduce the number of sites that can be surveyed (Guillera-Arroita and Lahoz-Monfort 2012). The challenge facing managers is knowing – ahead of time – if sufficient resources are being allocated to confidently detect population trends, or if resources are being spent in a way that maximises the chance of trend detection (Rhodes *et al*. 2006; Bailey *et al*. 2007).

Power analysis is a useful, yet under-utilised tool, to inform the design and/or assessment of wildlife monitoring programs. Statistical power of a monitoring design is defined as the probability that a null hypothesis of no change of interest will be rejected if such a change has truly occurred (Steidl *et al*. 1997; Strayer 1999). Power calculations require specification of: the change of interest that the designer wishes to detect, also known as the effect size (e.g. a 10% change in abundance over 10 years); acceptable type 1 error rate (false alarm rate – traditionally set at 0.05); the sample size of the monitoring program (e.g. number of sites being surveyed), and an estimate of 'natural' or background variation in the observed data. Background variation comprises stochastic environmental variation and observation error (e.g. counting error or detection error). Although statistical power can be directly calculated for very simple problems, in the practice of monitoring design, it generally involves simulating future observation data, taking into account natural and observation variation (ideally estimated by examining pilot data) and assumptions about how populations will change (the effect size). A statistical test is then conducted on the simulated data. This process is repeated numerous times to evaluate the proportion of times that the false null hypothesis of no (or little) change is rejected. This proportion is the statistical power of a given monitoring program.

This chapter describes how a power analysis has been used to review and re-design a long-running terrestrial vertebrate monitoring program in northern Australia. The monitoring program was initiated in the mid-1990s, and expanded in the early 2000s to track changes in the abundance and distribution of species, particularly small to medium-sized mammals, in the tropical savannas of northern Australia. The current monitoring program consists of ~300 sites located in three National Parks (Kakadu, Nitmiluk and Litchfield). Sites are surveyed every 5 years, applying a standard methodology (see overview in Woinarski *et al*. 2010); these data have been crucial for detecting dramatic population declines, as well as better understanding the role of drivers such as fire on species and communities (e.g. Woinarski *et al*. 2010; Woinarski *et al*. 2012; Lawes *et al*. 2015).

Although the monitoring program was instrumental in documenting population trends, an evaluation of its current design was needed for several reasons. First, the location of sites and frequency of sampling was not chosen with vertebrates in mind; rather, the program was super-imposed over an existing monitoring program designed to learn about the effect of fire on vegetation communities (Russell-Smith *et al*. 2014). Second, the densities of many species have declined so markedly since the program's inception that the original design now fails to detect many species at sufficient frequencies to allow future assessment of population trends. Third, there is concern that a 5-year survey frequency may not provide enough forewarning to trigger timely management in response to further population declines (e.g. Yoccoz *et al*. 2001).

This chapter describes a power analysis by simulation to help evaluate and re-design this long-term monitoring program. To do this, the occupancy and detectability of species recorded during past surveys was modelled to better understand the relationships between species distributions and environmental variables. Then declines in occupancy over time for nine species were simulated and the power of competing monitoring designs to detect simulated changes in occupancy was estimated. This analysis is 'spatially explicit' because it predicted occupancy and detectability of the nine species across six parks in northern Australia, not just at sites where monitoring data are collected. Some preliminary results demonstrate how this analysis will be used to inform the design and evaluation of long-term monitoring programs in the Northern Territory Parks.

Method

Determining the current state and detectability of selected populations

The first step in the analysis was to model occupancy and detectability for a selection of species recorded during surveys. Occupancy is defined as the proportion of sites, patches, or habitat units occupied by a species, while detectability is the probability of finding a species at a site if it is present (MacKenzie and Royle 2005). Both occupancy and detectability can depend on environmental variables, such as terrain or vegetation type.

Occupancy and detectability were modelled for nine of the 247 species that have been recorded during monitoring (Table 20.1). The selected species covered three taxonomic groups (birds, mammals and reptiles) and varied according to rarity and detectability.

Occupancy and detectability models were built for these nine species using detection/non-detection data collected over consecutive nights at ~300 sites in Kakadu, Nitmiluk and Litchfield National Parks, as well as sites in several other parks and reserves where one-off surveys have occurred (Fig. 20.1). Importantly, only the most recent data were used to fit models (collected from 2011–2015) to provide an up-to-date 'snapshot' of occupancy and detectability for the nine species.

For each species, it was assumed that occupancy could depend on a set of site covariates: elevation; terrain roughness; soil moisture; maximum temperature at the driest period; fire frequency; time since fire; and distance to perennial creeks. It was assumed that detectability could depend on time since fire, fire frequency, terrain roughness and the type of detection method (i.e. live trapping *v.* camera trapping). Occupancy and detectability maps were generated for six of the largest National Parks in the Northern Territory – Garig Gunak Barlu, Gregory, Kakadu, Limmen, Litchfield and Nitmiluk – with values ranging from 0–1 (see Fig. 20.1).

Table 20.1. Nine species selected for the spatially explicit power analysis with mean estimates of occupancy and detectability (for one day/night of surveying) across the landscape.

Group	Common name	Scientific name	Occupancy	Detectability
Birds	Partridge pigeon	*Geophaps smithii smithii*	0.04	0.29
	Rufous whistler	*Pachycephala rufiventris*	0.31	0.47
	Bar-shouldered dove	*Geopelia humeralis*	0.47	0.54
Mammals	Common planigale	*Planigale maculata*	0.15	0.07
	Common rock-rat	*Zyzomys argurus*	0.16	0.30
	Northern brown bandicoot	*Isoodon macrourus*	0.21	0.85
Reptiles	Northern spotted rock dtella	*Gehyra nana*	0.16	0.13
	Port Essington ctenotus	*Ctenotus essingtonii*	0.21	0.37
	Bynoe's gecko	*Heteronotia binoei*	0.48	0.25

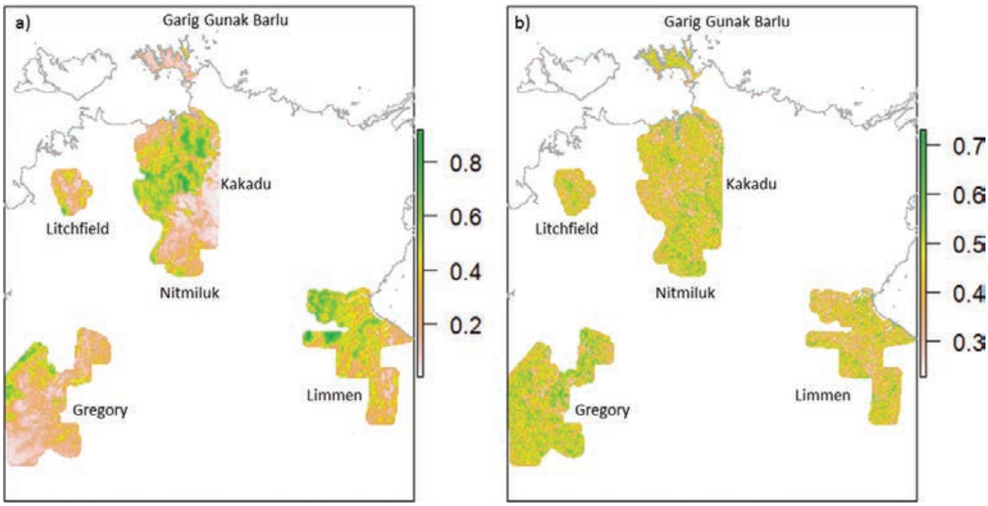

Fig. 20.1. Example map of predicted (a) occupancy and (b) detectability for 1 day/night of surveying for one of the nine species (rufous whistler *Pachycephala rufiventris*) across six parks in the Northern Territory. Data used to fit models were collected at ~300 sites sampled from 2011 to 2015.

The current occupancy status of each species was simulated in each cell of the occupancy maps. To determine which cells in the landscape are currently occupied, a 1 (present) or 0 (absent) was assigned for each species by comparing a random draw from a uniform (0,1) distribution with the occupancy probability for that cell. Draws less than the probability of occupancy resulted in a presence, and a 1 in the landscape for that species, otherwise cells were assigned a value of 0 (absent).

Simulating a trend in occupancy over time

A decline in occupancy was simulated for each species during each year of a monitoring program. The magnitude of the change in occupancy between the start and end of the monitoring program is hereafter referred to as the 'effect size'. For example, if the initial occupancy of a cell for a species was 0.8 at the start of a monitoring program and the effect size was 0.5, occupancy reduced to 0.4 at the end of the monitoring period. The sampling procedure described above was repeated at each point in time to determine the occupancy status of cells in response to a constant decline in occupancy over time. It was not known what the effect size will be, so a range of values from 0.1–0.9 were simulated over a 15-year period.

Simulating detection histories at monitoring sites

After simulating declines in occupancy, the process of data collection was simulated by generating 'detection histories' for each species at monitoring sites with each round of monitoring. To construct detection histories, a 1 (detection) or 0 (non-detection) was assigned for each visit to each occupied cell by comparing a random draw from a uniform (0,1) distribution with the detection probability of the particular species. Draws less than the probability of detection resulted in a detection of a species and a 1 in the detection history for that survey day and year. Draws greater than the probability of detection resulted in a 0 (non-detection). For example, a detection history of [0,1,0,0] indicated that a site was surveyed for four nights, and that the species was detected on only the second night.

Detecting trends in simulated data

After generating detection histories, simulated data were analysed to determine if there was a trend in occupancy over years. A two-tailed test was conducted and a trend was assumed as significant if the 95% confidence interval around the trend parameter excluded 0 (i.e. assuming a significance level of 0.05). Simulations of occupancy and data collection were repeated 1000 times, each time testing for a significant trend in occupancy over time. Statistical power was calculated as the proportion of simulations in which a significant trend was detected. This process was repeated for each species, at a range of effect sizes under alternative monitoring scenarios to enable comparisons of statistical power.

Alternative monitoring scenarios

Statistical power was estimated for three alternative monitoring designs: (1) 250 sites are surveyed every 5 years; (2) 150 sites are surveyed every 3 years; and (3) 107 sites are surveyed every 2 years. For all scenarios, it was assumed that sites were surveyed for four consecutive nights. The relationship between the frequency of surveys and the number of sites was determined by a cost model not described here.

Results

Preliminary results suggest that the ability to detect further population declines of the nine species is maximised by surveying sites every 3–5 years (Fig. 20.2). Due to budget constraints, surveying sites more regularly decreased the number of re-visits to sites that could be achieved, resulting in a reduction in detection rates and a commensurate reduction in statistical power. This highlights the essential trade-off between sites and visits that is difficult to get right without explicit analyses such as those presented here.

Statistical power of around 0.8 could be achieved for moderate-to-large declines in the most common and easily detectable species across the six parks (Fig. 20.2). Statistical power to detect declines was low for species with moderate initial occupancy and detectability, unless simulated declines were very large (i.e.

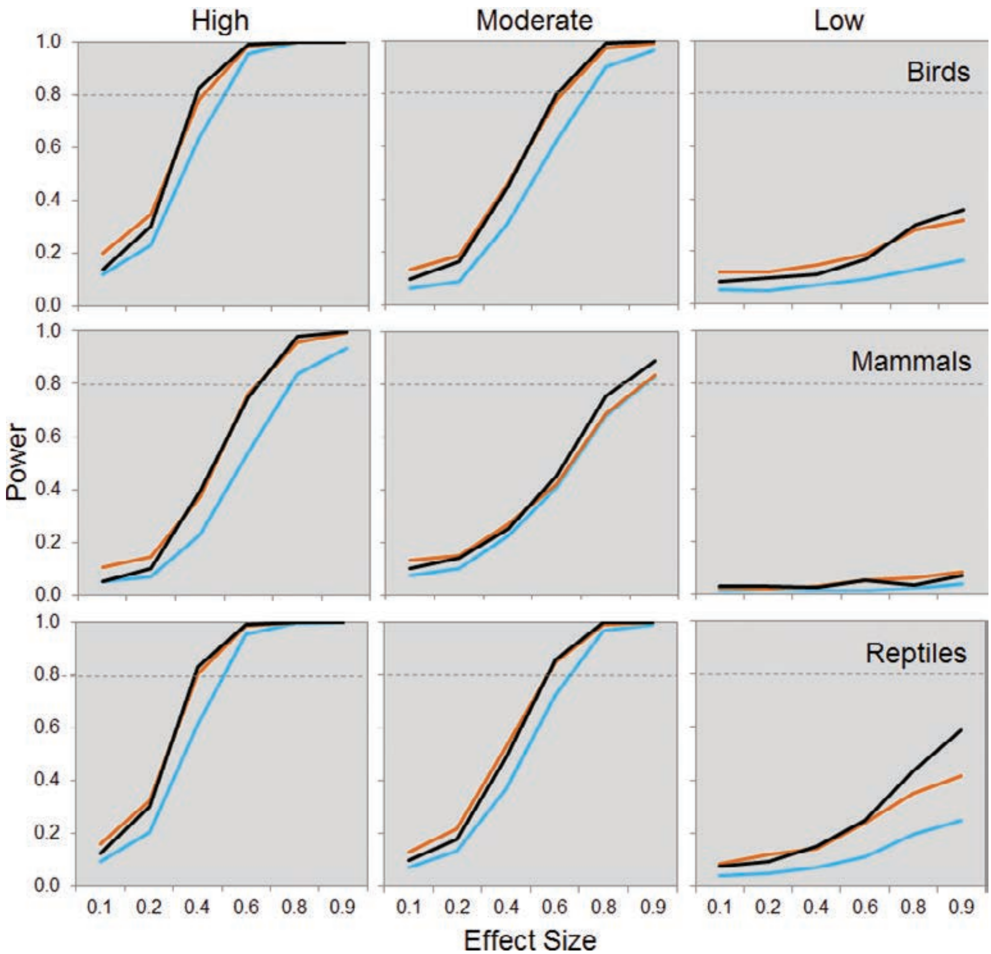

Fig. 20.2. Statistical power (*y*-axis) to detect declines of various effect sizes (*x*-axis) for three birds (top row), three mammals (middle row) and three reptiles (bottom row) with high (left column), moderate (middle column), and low initial (right column) initial occupancy/detectability for three monitoring designs: 107 sites surveyed every 2 years (blue line); 150 sites surveyed every 3 years (orange line); and, 250 sites surveyed every 5 years (black line). The dashed line indicates 80% power.

an effect size of 0.6–0.8 over 15 years; Fig. 20.2). For all species with the lowest initial occupancy and detectability, monitoring was unable to detect even severe declines (Fig. 20.2).

Discussion

Preliminary results reveal that detecting even large trends (50% decline in 15 years) in the majority of species across the six parks in the Northern Territory would require large-scale intensive sampling beyond the capacity of the current terrestrial vertebrate monitoring program. Despite this, the optimal number of sites and frequency of survey to maximise trend detection with the resources available was identified. The results presented here, combined with results of a broader analysis, have been used by managers to direct decisions in the re-design of the terrestrial vertebrate monitoring program in the Northern Territory.

The benefit of a spatially explicit power analysis is that potential survey locations can be positioned at any location in the landscape, enabling explicit evaluation of the efficiency of various designs taking into account travel costs, and the capacity to achieve adequate landscape-level stratification of sites across environmental gradients and management units. This opens up the possibility of landscape-level optimisation of sampling designs to detect overall, or management-related changes in abundance and occupancy.

This chapter has demonstrated how monitoring can be expanded into areas where we do not currently have data on species of interest. This approach allows assessment of a wide range of potential monitoring designs (Ellis *et al*. 2014; Steenweg *et al*. 2016), and exploration of other trade-offs such as the effect of placing sites in areas of highest or lowest occupancy and detectability; or the trade-off between monitoring frequency, the number of repeat visits to sites in each year or season, the method of detection employed, the number of species monitored, taking into account the costs of each monitoring design decision.

The power analysis required pilot data to estimate occupancy and detectability now and over time. When initial trapping data, home range or habitat use data are not available, expert opinion on habitat preferences, detection rates and initial distributions must be used in the form of habitat suitability indices (Burgman *et al*. 2001). However, it is emphasised that simulations become more realistic for species with accurate occupancy and detectability models and maps, preferably based on strong biological survey data.

Effective monitoring must be sensitive enough to detect trends with sufficient confidence when trend detection is the primary purpose of monitoring (Possingham *et al*. 2011). Many monitoring programs of threatened vertebrates profess to monitoring trends, but power analyses are rarely conducted when designing monitoring programs, so little knowledge of capacity to meet intended objectives exists. The application of power analysis is recommended where managers and/or researchers are attempting to measure change.

Power analysis is a useful tool for designing and re-evaluating threatened species monitoring programs because it can inform: how likely a new or existing monitoring design will achieve an objective; how resources can best be allocated to maximise the chance of meeting an objective; and how much additional funding is needed to achieve a sufficient level of statistical power. This information can justify existing monitoring programs, identify when further resources are needed, test the performance of new or alternative detection methods, and identify when monitoring is likely to waste resources that could otherwise be spent on management.

Lessons learned

- Simulation is useful for predicting how ecological systems might change over time while accounting for stochasticity.
- Power analysis can be a simulation tool for assessing – ahead of time – how likely a monitoring program will detect trends or changes in populations.
- A power analysis can assess the performance of an existing/new monitoring program, or be used to compare alternatives where the number of sites, frequency of visits, and survey duration differ.
- Power analysis is under-utilised and not many managers apply it in the design of monitoring programs.
- Even long-running and relatively well-funded monitoring programs can have low power – as was the case for the case study presented here.

References

Bailey LL, Hines JE, Nichols JD, MacKenzie DI (2007) Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. *Ecological Applications* **17**, 281–290. doi:10.1890/1051-0761(2007)017[0281:SDTIOS]2.0.CO;2

Burgman MA, Breininger DR, Duncan BW, Ferson S (2001) Setting reliability bounds on habitat suitability indices. *Ecological Applications* **11**, 70–78. doi:10.1890/1051-0761(2001)011[0070:SRBOHS]2.0.CO;2

Ellis MM, Ivan JS, Schwartz MK (2014) Spatially explicit power analyses for occupancy-based monitoring of wolverine in the U.S. Rocky Mountains. *Conservation Biology* **28**, 52–62. doi:10.1111/cobi.12139

Guillera-Arroita G, Lahoz-Monfort JJ (2012) Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods in Ecology and Evolution* **3**, 860–869. doi:10.1111/j.2041-210X.2012.00225.x

Lawes MJ, Murphy BP, Fisher A, Woinarski JCZ, Edwards AC, Russell-Smith J (2015) Small mammals decline with increasing fire extent in northern Australia: evidence from long-term monitoring in Kakadu National Park. *International Journal of Wildland Fire* **24**, 712–722. doi:10.1071/WF14163

MacKenzie DI, Royle JA (2005) Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology* **42**, 1105–1114. doi:10.1111/j.1365-2664.2005.01098.x

Possingham HP, Fuller RA, Joseph LN, Wintle BA (2011) Choosing among long-term ecological monitoring programs and knowing when to stop. In *Design and Analysis of Long-term Ecological Monitoring Studies*. (Eds R Gitzen, J Millspaugh, A Cooper and D Licht) pp. 498–508. Cambridge University Press, Cambridge, UK.

Rhodes JR, Tyre AJ, Jonzen N, McAlpine CA, Possingham HP (2006) Optimising presence absence surveys for detecting population trends. *The Journal of Wildlife Management* **70**, 8–18. doi:10.2193/0022-541X(2006)70[8:OPSFDP]2.0.CO;2

Russell-Smith J, Edwards AC, Woinarski JCZ, Fisher A, Murphy BP, Lawes MJ, *et al.* (2014) North Australian tropical savannas: the Three Parks Savanna Fire-Effects Plot Network. In *Biodiversity and Environmental Change: Monitoring, Challenges and Direction.* (Eds D Lindenmayer, E Burns, N Thurgate and A Lowe) pp. 335–378. CSIRO Publishing, Melbourne.

Steenweg R, Whittington J, Hebblewhite M, Forshner A, Johnston B, Petersen D, *et al.* (2016) Camera-based occupancy monitoring at large scales: power to detect trends in grizzly bears across the Canadian Rockies. *Biological Conservation* **201**, 192–200. doi:10.1016/j.biocon.2016.06.020

Steidl RJ, Hayes JP, Schauber E (1997) Statistical power analysis in wildlife research. *The Journal of Wildlife Management* **61**, 270–279. doi:10.2307/3802582

Strayer DL (1999) Statistical power of presence-absence data to detect population declines. *Conservation Biology* **13**, 1034–1038. doi:10.1046/j.1523-1739.1999.98143.x

Woinarski JCZ, Armstrong M, Brennan K, Fisher A, Griffiths AD, Hill B, *et al.* (2010) Monitoring indicates rapid and severe decline of native small mammals in Kakadu National Park, Northern Australia. *Wildlife Research* **37**, 116–126. doi:10.1071/WR09125

Woinarski JCZ, Fisher A, Armstrong M, Brennan K, Griffiths AD, Hill B, *et al.* (2012) Monitoring indicates greater resilience for birds than for mammals in Kakadu National Park, northern Australia. *Wildlife Research* **39**, 397–407. doi:10.1071/WR11213

Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution* **16**, 446–453. doi:10.1016/S0169-5347(01)02205-4