

Article

Adaptive monitoring using the endangered northern corroboree frog (*Pseudophryne pengilleyi*) as a case study

Francis Lemckert^{1,2}, Trent Penman¹, Andrew Haywood^{1,3}

¹Forest Science Centre, NSW Department of Primary Industries, PO Box 100, Beecroft, NSW. 2119. Australia

²School of Biological and Earth Sciences, The University of Newcastle. University Drive, Callaghan, NSW. 2380. Australia

³Current Address: Department of Sustainability and Environment, 3/8 Nicholson St, East Melbourne, VIC. 3002. Australia

E-mail: frankl@sf.nsw.gov.au

Received 11 March 2011; Accepted 18 April 2011; Published online 28 August 2011

IAEES

Abstract

Monitoring programs are most successful when they undertake regular evaluation of their data to determine if the goals of the programs are achievable and allow changes to achieve this as necessary – so called adaptive monitoring. We use data from a monitoring program for the northern corroboree frog (*Pseudophryne pengilleyi*), a declining species in south-eastern Australia, to determine the inter-annual variability in the counts and assess what levels of population change would be detectable using different statistical and monitoring approaches. The existing monitoring program would only successfully statistically detect a 3% annual decline (34% total decline) in population size over a ten year period. Monitoring 40 sites would allow an 80% or greater chance of detecting a 2% or greater annual increase over a ten year period (22% increase). Detecting population decreases is more difficult as monitoring 40 sites with a 2% annual decline (19% total decline) will have a less than 40% chance of being detected after 10 years. A larger monitoring program is required to detect smaller annual changes in the population of this species. These findings have implications for the likely effectiveness of other anuran monitoring programs as the northern corroboree frog appears to be far more consistent in detectable call effort compared to most species.

Keywords adaptive monitoring; *Pseudophryne pengilleyi*.

1 Introduction

Monitoring populations of animals is an important means of assisting in their conservation and are used widely (e.g., Alien et al., 1996; Kirk and Hyslop, 1998). Effective programs can alert managers to declines that may threaten the long-term status of a population or increases that indicate success in implemented strategies (e.g., Noss and Cooperrider, 1994; Funk et al., 2003). Many programs however, are undertaken on an ad hoc basis, with counts being undertaken to monitor if populations are remaining stable, but without an understanding if an analysis of that data would be able to effectively detect a change in the population or even what change would be considered important to detect (Witmer, 2005; Field et al., 2007). An important step in any such situation is to review the data after a suitable period of time to determine its capability to ascertain population changes. Changes to the program can then be implemented if required, to ensure that the program can achieve its goals (Field et al., 2007). Such adaptive management appears to be rarely undertaken (Field et al., 2007), possibly because the value of this approach is not fully appreciated.

Anuran amphibians present a particularly difficult group to monitor (Marsh 2001). Adults are not readily censused in most environments, due to their cryptic habits, and so counts are usually based on breeding choruses of male frogs (e.g., Crouch and Paton, 2002; Buckley and Beebee, 2004; Lemckert et al., 2004). These can be readily undertaken, but are strongly weather dependent and so highly variable between weeks, days and sometimes even hours (e.g., Einem and Ober, 1956; Blankenhorn, 1972; Robertson, 1986; Krupa, 1994; Bridges and Dorcas, 2000; Lemckert, 2001). This greatly reduces their sensitivity when applied to monitoring overall population sizes. Given the widespread nature of this type of monitoring in anurans, determining if this approach can be used accurately in population monitoring would appear to be desirable. Careful estimation of count parameters may still allow them to be used for trend assessment, but this has been rarely tested.

For any species, it is impossible to measure the total population and therefore the analysis and subsequent conclusions are based on only a sample of the total population. In doing so, there is always a risk of making an incorrect inference, and two types of errors can arise:

Type 1: We conclude that there is a change in the population when, in reality, there is no change

Type 2: We conclude that there is no change when there is, in fact, an increase or decrease

Tests of statistical significance are designed to reduce the occurrence of Type 1 errors to an acceptable level (usually 5%). Only when this likelihood is smaller than 5% do we reject the “no change” hypothesis and accept that change has occurred. Power is the probability of rejecting the “no change” hypothesis when it is false. A study with low power will probably conclude there is no change in a population when a change does actually occur. Clearly such a monitoring program is a waste of resources and a threat to the conservation of a species, so we need to design monitoring programs with high power and halt or modify those that only have low power to detect change.

The northern corroboree frog (*Pseudophryne pengilleyi* Wells and Wellington) is a small myobatrachid frog found in southern highlands of New South Wales (NSW) and the Australian Capital Territory (ACT). It is listed as endangered in NSW and Vulnerable under Commonwealth legislation due to significant declines in its range and numbers since the 1970s (Osborne et al., 1999). A large population of northern corroboree frogs remains in Buccleuch State Forest of southern NSW, which is subjected to logging, road construction and burning. Forests NSW, the agency managing the area, commenced an annual monitoring program in the late-1990s to count calling males at a selected subset of the available breeding sites. The initial intent was to provide indications of declines in population within the forest, although the option for collecting data to test the influences of differing management strategies has also been considered. However, the potential to statistically detect changes in the population using this data has not been tested.

In this study, we examine a monitoring program’s ability to detect both increases and decreases in the population counts of the northern corroboree frog. We consider the “power” of the current monitoring program and provide information on what levels of monitoring are required to meet varying monitoring outcomes and the most appropriate tests to achieve these outcomes.

2 Methods

Northern corroboree frogs breed in montane, swampy areas above 900 m ASL. Males call for only around three weeks each year, starting some time between mid-January and late February, but mostly in February (Osborne, 1989). Hence there is a very limited calling and breeding season for this species. Eggs are deposited in nests situated in areas that are inundated at a later time. The tadpoles hatch at an advanced stage when the nest chamber is flooded and develop in pools in the swamp to metamorphose in late spring to mid-summer.

During non-breeding periods, adults and juveniles appear to spend the majority of time in woodlands surrounding the breeding sites (Osborne et al., 1990).

The northern corroboree frog population evaluated in this study occurs within and around Buccleuch State Forest, which is centred approximately 35 km north of Tumut, on the western side of the southern highlands of New South Wales (Fig. 1). This forest covers an area of 64,630 hectares, 46,700 of which have been converted to plantations (mainly *Pinus radiata*). The remainder is native forest comprised mainly of a mixed Broad-leaved Peppermint (*Eucalyptus dives*) and Manna Gum (*E. viminalis*) sclerophyll forest (7000 ha), but also with significant stands of Snow Gum (*E. pauciflora*), Manna Gum and Alpine Ash (*E. delegatensis*) forest (D. Leslie Pers. Comm.). The altitude ranges from 780m to 1440 m, with a mean annual rainfall of 1350 mm and mean monthly minimum and maximum temperatures of 4° C and 5° C, respectively. The area has been logged on a rotational basis since the early 1900s, generally using selective logging practices. Plantation establishment commenced in the 1920s, but major acquisition of freehold or leasehold lands and subsequent conversion to plantation occurred in the 1950s.

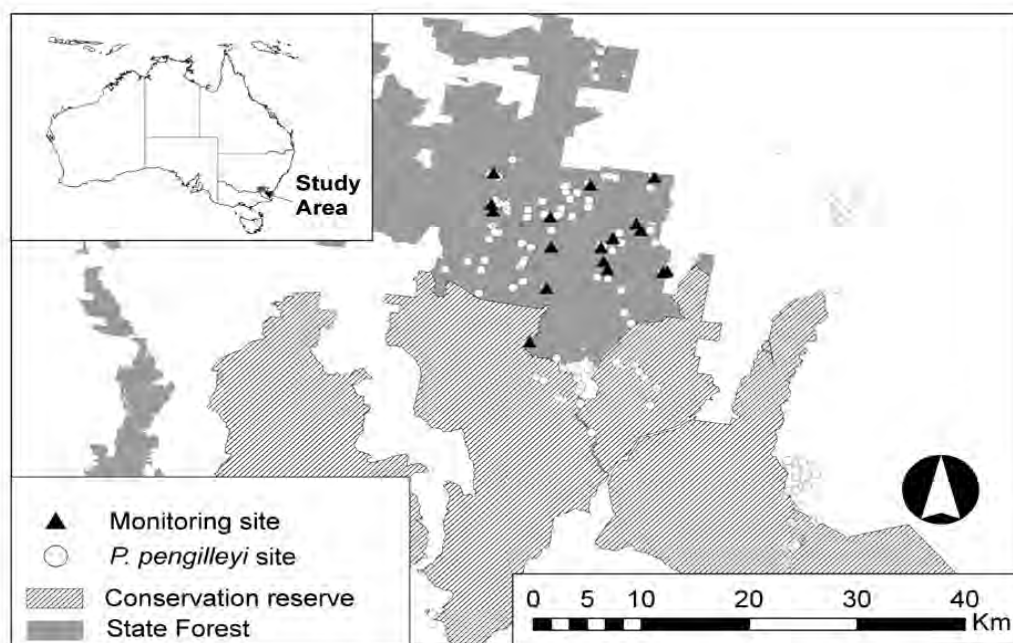


Fig. 1 Study area with locations of monitoring sites and other sites where the northern corroboree frog has been recorded from

3 Monitoring

Count data numbers of calling males have been obtained by Forests NSW staff for up to 18 sites in 1997 then annually from 1999 through to 2006 (Table 1). Preliminary monitoring was conducted to determine when the species commenced calling. Counts were undertaken in the second of the three week calling period which is considered the high point of calling. The counts were undertaken between 8:00 and 18:00 hrs by generally two to four people driving from one site to the next in a fixed order over a two-day period. The extent of the swamp area searched was approximately the same each year and observer error was reduced as far as possible by having two or more people involved in each count to provide an average score. At each site, the surveyors walked to the edge of the designated area and spent an initial one to two minutes listening to count the numbers of calling male corroboree frogs. The site was then systematically traversed with the surveyors using

loud vocalisations (usually shouting “Oiy” or “Hello Frogs”) eliciting territorial responses of resident males. This is a highly effective means of obtaining consistent responses by male frogs and individual males are easily identified (F. Lemckert, Pers. Obs.). The numbers of calling/responding males were recorded, along with the time, date and weather conditions. The number of males located is used as a surrogate for population size, with the expectation that any changes in total population size will be reflected proportionally in the numbers of adult males present and calling at the sites.

Table 1 Count values for 18 sites searched at varying frequencies between 1997 and 2006

Site	1997	1999	2000	2001	2002	2003	2004	2005	2006	MEAN	STD	CV
1	6	2	5	-	1	3	7	9	6	4.9	2.7	0.6
2	16	3	0	-	-	3	-	-	-	5.5	7.1	1.3
3	16	4	7	12	6	4	0	1	6	6.2	5.1	0.8
4	3	0	0	0	0	0	0	0	-	0.4	1.1	2.8
5	34	34	50	37	35	15	13	50	50	35.3	14	0.4
6	30	25	20	8	4	3	7	12	12	13.4	9.5	0.7
7	34	23	24	15	14	12	7	12	7	16.4	8.9	0.5
8	0	0	2	1	0	0	0	0	-	0.4	0.7	2
9	0	1	5	0	0	0	1	2	0	1	1.7	1.7
10	9	1	4	2	8	2	0	0	0	2.9	3.4	1.2
11	0	1	0	0	0	0	2	7	6	1.8	2.8	1.6
12	30	25	40	25	19	2	22	35	25	24.8	10.7	0.4
13	9	4	4	1	1	0	0	0	-	2.4	3.2	1.3
14	12	1	4	-	-	1	-	-	-	4.5	5.2	1.2
15	4	1	3	-	-	0	-	-	-	2	1.8	0.9
16	8	6	5	-	-	1	-	-	-	5	2.9	0.6
17	32	35	31	17	27	3	26	35	20	10.4	25.1	0.4
18	14	9	25	40	12	7	18	20	15	10	17.8	0.6

4 Design of Sensitivity Analysis

Before conducting the sensitivity analysis, we ran a random effects model to determine if there were any significant trends over time in the existing data set. Only sites with more than four years of survey data were used, leaving 14 available sites for the analysis. A random effects model found a significant decline across all sites over time (estimate = -0.699 ± 0.222). To create a stable population for the study period for the simulations all data were de-trended. The creation of a stable population was necessary to accurately test the ability of the program to detect real increases or decreases in population size.

The sensitivity analysis was designed to simultaneously affect the ability of a monitoring program to detect either increases or decreases in population trends over time with varying numbers of sites for a ten year period. We calculated the ability to detect annual changes in total population size of 0 to 10% at 1% increments for both increasing and decreasing trends. For example, a 5% decline means that each year the population is reduced by 5% of the size of the previous year. Over a ten year period this results in a total reduction of approximately 40% of its size at time 0. To test the sample size needed for an effective monitoring program, we simulated the number of sites in the monitoring program starting from 5 and increasing up to 100 in increments of 5.

Simulations were conducted using the following steps:

1. A random sample of n sites (where n is between 5 and 100) was taken from the original 14 sites with replacement. Using replacement, allows for the simulation sample size to exceed the number of original sites.

2. Calculate a mean and standard deviation from the detrended data set for each site in the random sample.
3. Calculate time 0 values for each site by taking a random number from a normal distribution with the mean and standard deviation from step 2.
4. Then for years 1 to 10, the mean was calculated by decreasing or increasing the original mean by the trend. For example, if the time 0 mean for a site was 10 and we were simulating a 10% decline, the time 1 mean would be 9 and the time 2 mean would be 8.1.
5. Values for each site were then calculated by taking a random number from a normal distribution with the mean for the year calculated in step 4 and standard deviation for the site calculated in step 2.
6. A total dataset was created for the “n” sites sampled to be used in the statistical analysis.

The simulated dataset was tested to determine whether changes in population size could be detected using three techniques – generalised linear models (GLM), generalised estimating equations (GEE) and random effects models (REM). GLM represents a simple approach where a simple Gaussian regression is fitted to the data over time. GEE's extend GLM's to account for serial correlation in repeated measures studies (Liang and Zeger 1986). In this case, the repeated measure is on each of the sites considered in the monitoring. REM represents another means of accounting for repeated measures of sites in a regression framework (Gelman and Hill, 2007). For each model, we determined whether a significant trend could be detected. The process was then repeated 100 times for each trend and sample size combination resulting in a total of 42 000 simulations. All analyses were conducted using the R-package v2.5.0 (R-core Development Team, 2007) with a number of additional packages - regress (Clifford and McCullagh, 2006), reshape (Wickham, 2007) and geepack (Yan, 2002; Yan and Fine, 2004).

5 Results

When present at a site, the numbers of frogs recorded ranged from 1 to 50 with a mean count per site of 9.43 frogs with a standard deviation of 5.62 and a median of 5. The variation in numbers of males counted at a site between years could be great, although the overall population changed to a much lesser degree (Fig. 2). For example, the greatest yearly total of frogs was recorded in 1997 with 260 frogs from 21 sites and the lowest was 56 frogs from 20 sites in 2003, when a severe drought was in progress. In 2003, eight of 20 sites returned a zero count whilst, in 1999, only four of 22 sites had zero counts.

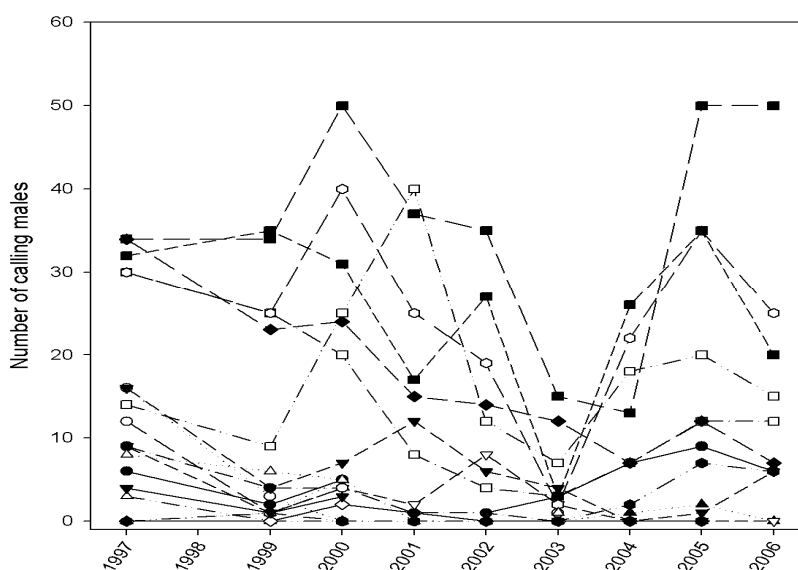


Fig. 2 Fluctuations in the recorded numbers of frogs for the 14 sites used in the simulation analysis

6 Sensitivity Analysis

The sensitivity analysis indicates that if the current monitoring strategy continues at similar levels (i.e. 14 sites), after ten years the program has sufficient power to detect a 6% annual increase (80% increase in the number of calling males) using a GLM or GEE approach or a 4% annual increase (48% increase in the number of calling males) using a REM approach (Fig. 3). In contrast, the program has sufficient power to detect a 7% annual decrease (52% decrease in the number of calling males) using GLM or GEE's or a 4% annual decrease (34% decrease in the number of calling males) using a REM approach (Fig. 4).

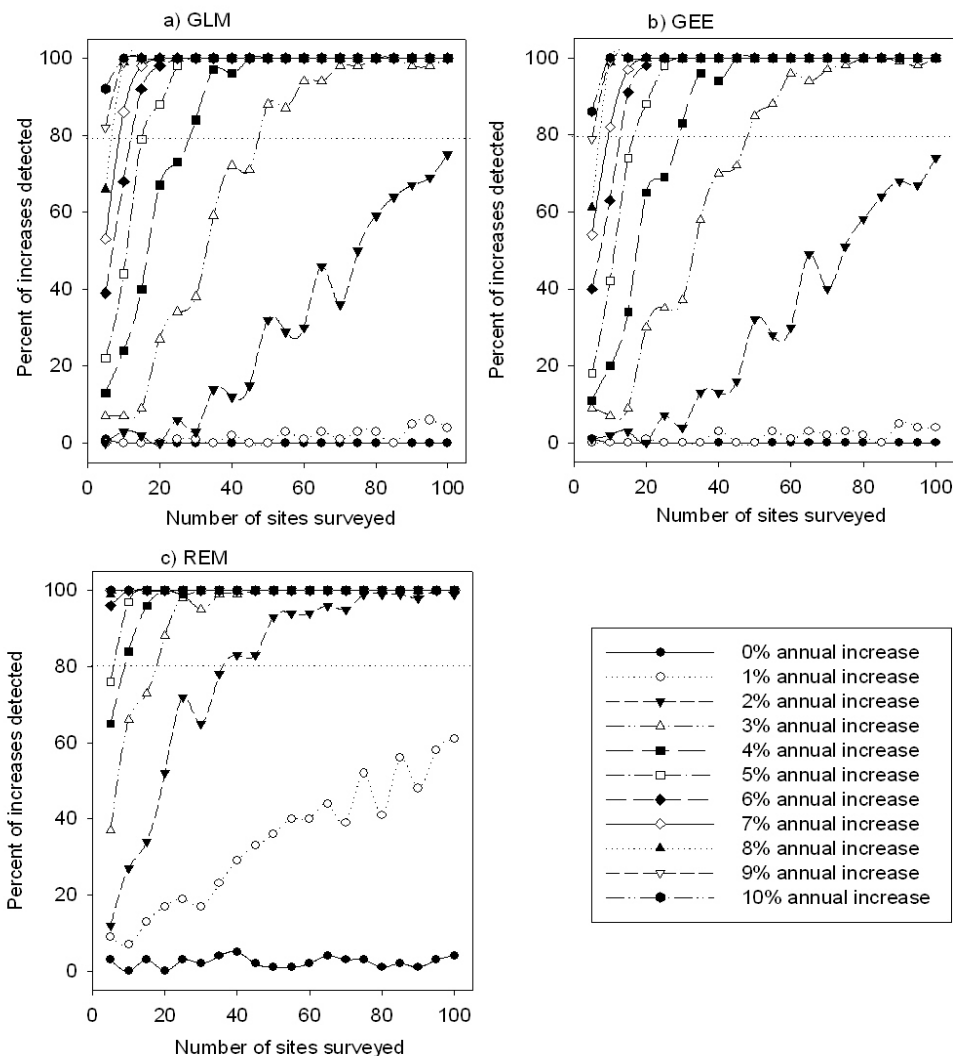


Fig. 3 Probabilities of detecting differing levels of annual increase in the northern corroboree frog populations in the hardwood forests near Tumut, NSW. Nb. The vertical line represents the probabilities of detecting changes for 14 monitoring sites.

For all analyses, increasing trends were more readily detected than decreasing trends and the REM approach was far more sensitive than both the GEE and GLM approaches. There were only minor differences between the sensitivity of the GLM and the GEE approaches, with neither being consistently more sensitive (Figs 3 and 4). Both the GLM and GEE approaches failed to gain sufficient power to detect a 2% annual increase (21% increase in population size) even with 100 sites, whereas the REM could detect a 2% annual increase with

only 40 sites. In the decreasing scenarios, GLM and GEE also failed to detect a 4% annual decrease (36 % decrease in population size), but the REM could detect this change with 40 sites.

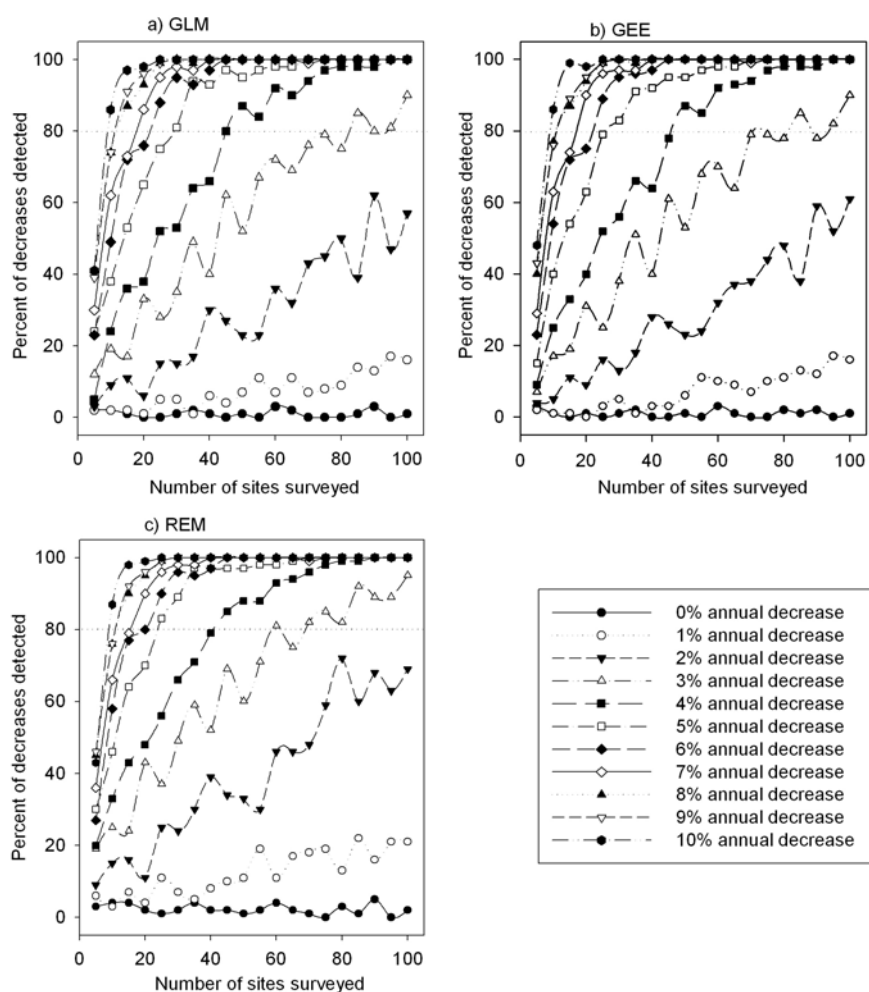


Fig. 4 Probabilities of detecting differing levels of annual decrease in the northern corroboree frog populations in the hardwood forests near Tumut, NSW. Nb. The vertical line represents the probabilities of detecting changes for 14 monitoring sites.

7 Discussion

Simulating changes in the measures of activity provides important information for the development of future monitoring programs for this species. Our testing of the data using standard analyses indicates that the existing level of survey effort will only statistically detect catastrophic decreases or population explosions. More sensitive statistical methods, such as random effects models, may alleviate part of this problem however increased replication of monitoring sites remains the most effective means of ensuring that populations are managed effectively. This work highlights the importance of understanding the power of any monitoring program being developed.

The results of the simulations enable us to develop an improved monitoring program that should be able to detect a decline in the overall population estimate obtained. We need to make two decisions to do this. Firstly, we need to select a level of decline in the population estimate that is considered significant enough to implement conservation management actions to reverse the decline. We do not yet have a good understanding

of natural fluctuations and population changes. However, the more important question is what represents a critical reduction in population size and over what time frame? For example, an annual change of 2.5% in population size may not seem large, but over a ten year period, this results in almost a 25% reduction in the total population size. In an adaptive management framework, this value could be varied as more information became available. Secondly, we need to consider the level of confidence or power we require in our program. Given that there is natural variation in the population estimates each year and survey results are variable, there is some possibility that we will detect a change that is not really a significant change and some chance that we could miss a significant decline in the population estimate. The choice of statistical test applied can affect this. In this study, the REM's detected changes more readily than GLM or GEE at the cost of detecting a small number of changes in the 0% category. REM's account for within and between site variability whereas the GLM and GEE approach average across the entire range of sites tested (Wolfinger and O'Connell, 1993). In the conservation management of threatened species, we are most concerned with responding to declines and therefore would adopt a more sensitive approach.

7.1 Future options for the program

The monitoring data on which we have based our assessments is relatively limited. Hence, our estimates of the natural variation in calling activity, and so population size, have limited certainty. Our analysis indicates that a minimum of 40 sites be searched annually in order to achieve a suitable level of power to detect changes in this population. Numerous other sites exist for expansion of the monitoring program (see Fig. 1) and this would provide the additional data to increase our ability to accurately estimate population levels and changes. Whilst Forests NSW could increase the number of sites being monitored to achieve this, there are two other monitoring programs for this species being conducted in this same area that could be combined with this to achieve the number of sites required for a suitable program. The NSW DECC and plantations division of Forests NSW both also monitor this species on the lands under their respective jurisdictions and a properly coordinated count program would enable the data to be combined. This requires each program to use the same methods (they essentially already do), with proportionally equal time allocated to detecting males at each site to maintain a consistent level of effort. Sampling would also need to be carried out over the same time period of each calling season. The start of the northern corroboree frog breeding season does vary from year to year, depending on environmental conditions. It is believed that the peak of calling occurs in the second of the three week annual calling season, rising and falling either side of this (D. Leslie Pers. Comm.).

Whatever the case, the data should be assessed after each season to determine if the power of the program is sufficient to detect the decrease in population size estimates set for the program and the chances we have in missing a significant change or detecting a non-significant change. Such an approach of adaptive monitoring is widely advocated (Green et al., 2005; McCarthy and Possingham, 2007). The number of sites being monitored each year can then be increased or decreased, if necessary, to ensure the program can detect significant population changes.

As the certainty in detecting changes is increased, the local land managers can take the opportunity to answer more specific questions regarding population change. The impacts of logging or controlled burns are two ongoing issues that may need to be addressed by research incorporated with the monitoring program. To do so would require selecting a series of replicate disturbed or "experimental" sites that would be compared to an equal number of undisturbed control sites.

7.2 Overall applicability of the program

A program that cannot detect changes in numbers until they have reached critical levels may very well result in irreversible or long-term damage to the population. Frogs represent a very serious challenge in this regard because counts of anurans can and do naturally vary greatly between weeks and even days, depending on

climatic conditions. Knowing how effectively a monitoring program can detect changes in frog populations clearly then allows the manager or researcher to explicitly state how certain they can be in using the results to manage the population in the long-term. The process of developing a monitoring program that we have detailed here is equally applicable to any other species of frog and can be used as the basis for any person to do so. It is simply a matter of obtaining enough survey records to be able to understand the expected variation in monitoring counts, which allows us to determine what number of sites are needed to detect a predetermined change. What change, if any, is acceptable is a matter for land managers to determine weighing up the value of a land use against the impact it has on the environment.

Acknowledgments

We would like to acknowledge the work of David Leslie and Bob Gay in developing the monitoring program for the frog in Buccleuch State Forest and undertaking large parts of the fieldwork. Forests NSW Research Division has maintained an ongoing interest in managing wildlife populations on the forest estate and supported our work. Brad Law and David Hunter both provided comments on this work. This work has been carried out under appropriate Scientific and Animal Research Licences.

References

- Alien L, Engeman R, Krupa H. 1996. Evaluation of three relative abundance indices for assessing dingo populations. *Wildlife Research*, 23: 197-206
- Blankenhorn HJ. 1972. Meteorological variables affecting onset and duration of calling in *Hyla arborea* (L.) and *Bufo b. calamita* (Laur.). *Oecologia*, 9: 233-235
- Bridges AS, Dorcas ME. 2000. Temporal variation in anuran calling behaviour: Implications for surveys and monitoring. *Copeia* 2000(2): 587-592
- Buckley J, Beebe TJC. 2004. Monitoring the conservation status of an endangered amphibian: the natterjack toad *Bufo calamita* in Britain. *Animal Conservation*, 7: 221-228
- Clifford D, McCullagh P. 2006. *Regress: Gaussian linear models with linear covariance structure*. R package version 1.0-0. <http://galton.uchicago.edu/~clifford/>
- Crouch WB III, Paton PWC. 2002. Assessing the use of call surveys to monitor breeding Anurans in Rhode Island. *Journal of Herpetology*, 36: 185-192
- Einem GE, Ober LD. 1956. The seasonal behaviour of certain Floridian Salientia. *Herpetologica*, 12: 205-212
- Field SA, O'Connor PJ, Tyre AJ, et al. 2007. Making monitoring meaningful. *Austral Ecology*, 32: 485-491
- Funk CW, Almeida-Reinoso D, Nogales-Sornosa F, et al. 2003. Monitoring population trends of eleutherodactylus frogs. *Journal of Herpetology*, 37: 245-256
- Gelman A, Hill J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge, UK
- Green RE, Balmford A, Crane PR, et al. 2005. A framework for improved monitoring of biodiversity: Responses to the World Summit on Sustainable Development. *Conservation Biology*, 19: 56-65
- Kirk DA, Hyslop C. 1998. Population status and trends in Canadian raptors: a review. *Biological Conservation*, 83: 91-111
- Krupa JJ. 1993. Breeding biology of the Great Plains toad in Oklahoma. *Journal of Herpetology*, 28: 217-224
- Lemckert FL. 2001. The influence of micrometeorological factors on the calling activity of the Australian frog *Crinia signifera* (Anura: Myobatrachidae). *Australian Zoologist*, 31: 625-631
- Lemckert FL, Brassil T, Haywood A. 2004. Effects of low intensity fire on pond-breeding anurans in mid-northern New South Wales, Australia. *Applied Herpetology*, 1: 183-195

- Liang KY, Zeger SL. 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73: 13-22
- Marsh DM. 2001. Fluctuations in amphibian populations: a meta-analysis. *Biological Conservation*, 101: 327-335
- McCarthy MA, Possingham HP. 2007. Active adaptive management for conservation. *Conservation Biology*, 21: 956-963
- Noss RF, Cooperrider AY. 1994. *Saving Nature's Legacy: Protecting and Restoring Biodiversity*. Island Press, Covelo
- Osborne WS. 1989. Distribution, relative abundance and conservation status of Corroboree Frogs, *Pseudophryne corroboree* (Anura: Myobatrachidae). *Australian Wildlife Research*, 16: 537-547
- Osborne WS. 1990. The biology and management of the Corroboree Frog (*Pseudophryne corroboree*). NSW Species Management Report No. 8. New South Wales National Parks and Wildlife Service, Hurstville
- Osborne W, Hunter D, Hollis G. 1999. Population declines and range contraction in Australian alpine frogs. In: *Declines and Disappearances of Australian Frogs* (A Campbell ed). Environment Australia, Canberra, 145-157
- R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>
- Robertson JGM. 1986. Female choice, male strategies and the role of vocalization in the Australian frog *Uperoleia rugosa*. *Animal Behaviour*, 34: 773-784
- Wickham H. 2007. *Reshape: Flexibly reshape data*. R package version 0.7.4. R Foundation for Statistical Computing, Vienna, Austria
- Witmer GW. 2005. Wildlife population monitoring: some practical considerations. *Wildlife Research*, 32: 259-263
- Wolfinger R, O'Connell M. 1993. Generalized linear mixed models a pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*, 48: 233-243
- Yan J. 2002. Geepack: Yet Another Package for Generalized Estimating Equations. *R-News*, 2/3, 12-14
- Yan J, Fine JP. 2004. Estimating equations for association structures. *Statistics in Medicine*, 23: 859-880