MN346: Evaluation of biodiversity monitoring methods

Dr Peter Long, Oxford University

It is only worthwhile to commit resources to a biodiversity monitoring programme if the methods being used would permit a significant trend of a certain magnitude in an indicator (say, 10% decline) over a specified period (say, over 3 years) to be detected with a desired level of confidence (eg 95%).

Whether a particular method (eg forest bird point counts) will satisfy these criteria depends on the number of sample units, their spatial distribution, the number of sampling occasions on each unit each year, the prevalence in the landscape and detectability of the species of interest, the underlying variance in the state parameter estimated (ie. density, occupancy, relative abundance), and the level of disaggregation of the indicator (eg by individual species or by groups).

You could use our monitoring data from 2009, 2010, 2011 and 2012 and undertake a power analysis focused on one or more particular taxonomic groups eg wetland birds, forest birds, lemurs, reptiles and amphibians and contrasting the indicators derived from GLMM, distance sampling and occupancy models.

This approach is only possible because we have a very well organised data collection system and relational databases which take the hard work out of the data handling allowing you to focus on the analysis. However, first it's important to understand the way a few key terms are used in the Mahamavo monitoring programme.

Activity = one of the different sampling methods used in Mahamavo eg bird point counts, walking lemur routes.

Sample unit = place in the forest of known co-ordinates at which a scientific sampling activity takes place eg the named approx 3km long sample routes used for monitoring lemurs and herps. Also applies to point locations used for bird point counts.

Occasion = a particular time on which a scientific sampling activity takes place on a sample unit. Typically each sampling unit will be sampled on many occasions in a field season.

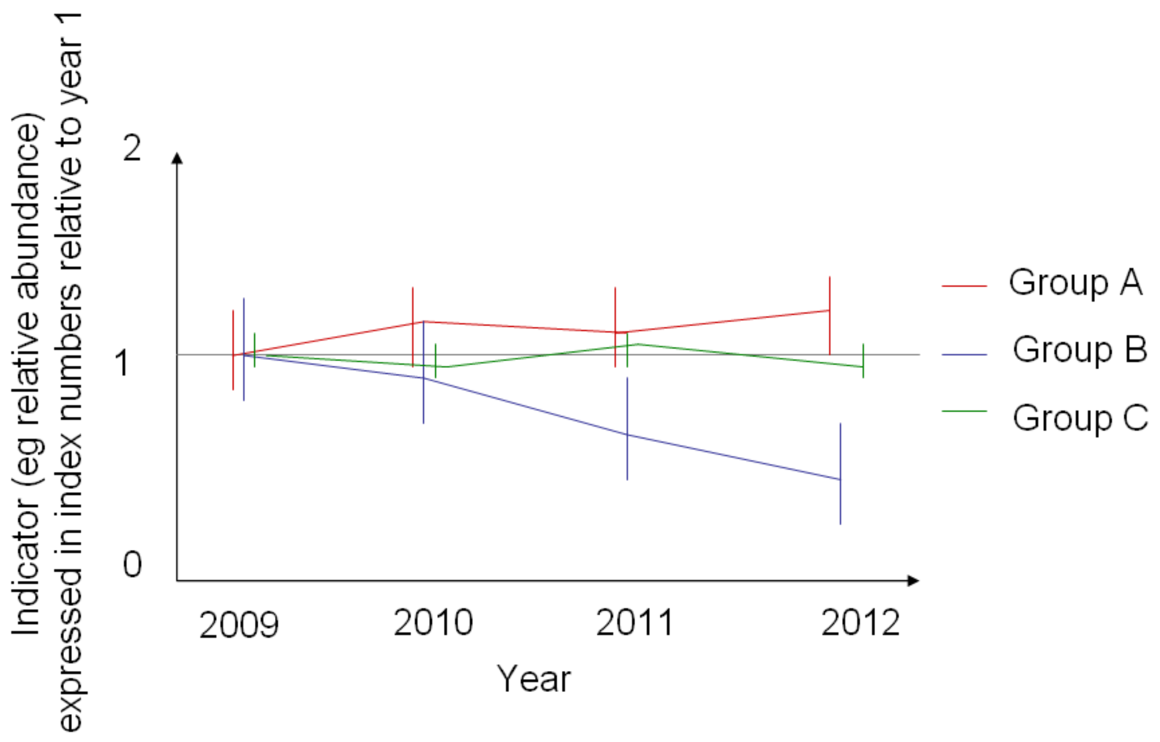
Record = an observation made on a specified occasion. On each occasion usually several records will be made, but it's possible that there might be no records on an occasion.

This creates a nested data hierarchy which can be manipulated in various ways. The most obvious things we do are to bind all the sample units with the occasions and then bind all the records with the occasions to create a very rich table which tells you the full context of every observation. It's also possible to work out things like detection histories, convert records into effort-corrected relative abundances over all occasions in that year (total number of individuals observed per unit effort eg minutes of point count or metres of route walked on that unit in that year). In turn these can then be cross-tabulated in various ways and disaggregated by species or by groups such as families or guilds (eg crocodiles, turtles, frogs, boas, skinks, chameleons, colubrid snakes, geckos). Here's an example of the structure of the results you might get from a database query on relative abundances.

Year	Sample unit	Group A	Group B	Group C	Group N
2009	1	0.0012			
2009	2	0.0000			
2009	3	0.0347			
2009	n	…			
2010	1				
2010	2				
2010	3				
2010	n				
2011	1				
2011	2				
2011	3				
2011	n				
2012	1				
2012	2				
2012	3				
2012	n				

Similar queries can produce perpendicular distances and cluster sizes for analysis in DISTANCE or detection histories for analysis using presence as an intermediate step to producing a similar table of a different kind of indicator.

The sketch graph shows an example of what a set of relative abundance indicators might look like for three hypothetical groups if you plotted mean RA across sample units by year (converted to index numbers). Note that if you estimate and plot standard errors within the set of units surveyed in a particular year you can quickly see if significant declines have occurred relative to the start of the monitoring programme by whether or not the error bar intersects the grey line. In this example Groups A and C have experienced no significant change whereas group B has declined.



However this approach neglects differences due to allocation of sampling effort to sample units in particular years. It is important to take account of this because we operate a rotating panel sampling design in which not all sample units are surveyed in on the same number of occasions in every year. This is a typical feature of almost all long-term biodiversity monitoring programmes. Since the sample units vary in their suitability for a given species, some variance in the indicator may be caused by the sampling effort allocation rather than by a real biological change. A better approach is to use GLMM, which is sometimes referred to as 'route regression' in the monitoring literature.

You could fit a generalised linear mixed model (GLMM), separately for each group of interest (using subset) of the form:

relative abundance ~ year | sample unit

This is easy to do using R libraries lme4 and/or MASS. It would then be possible to revise the graph with corrected error bars. This method can also be extended to any other indicator such as occupancy proportion from occupancy modeling of the sample units, total area of occupancy derived from time-explicit species distribution models, densities derived from distance sampling.

Now we come to the interesting research question: in what circumstances (indicator, taxonomic group, disaggregation level) has there been power to detect significant trend, from the Mahamavo database? To address this question would require performing multiple power analyses based on sample sizes and variances in contrasting groups (birds, lemurs, herps), indicators (relative abundance, occupancy, AOO, density) disaggregated in different ways (eg individual species, reptile groups, all reptiles) and then tabulating the number of years of monitoring necessary for a particular number of sample units or the number of sample units necessary for a fixed length of monitoring programme to detect a trend of a particular magnitude with a fixed level of confidence. This would allow trade-offs in monitoring design to be quantified and concrete recommendations to be made to improve the programme.

In principle this approach could also be extended to contrast different types of environments: tropical moist forest, tropical dry forest, savanna, by making comparisons between several OpWall sites eg Madagascar, South Africa, Honduras which is possible since the sampling methods and databases are almost identical between sites.

Further reading

Crawley MJ (2005) The R book. Wiley

Dixon PM, Pechmann JHK (2005) A statistical test to show negligible trend. Ecology 86: 1751-1756

Hoenig JM, Heisey DM (2001) The abuse of power: the pervasive fallacy of power calculations for data analysis. The American Statistician 55: 19-24

Manly FJ (2006) Randomisation, bootstrap and Monte Carlo methods in biology. Chapman and Hall

Loh J, Green RE, Ricketts T, Lamoreux J, Jenkins M, Kapos V, Randers J (2005) The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society Series B 360: 289-295

Quinn G, Keough M (2002) Experimental design and data analysis for biologists. Cambridge University Press.
See especially chapter 7 – design and power analysis

Royle JA, Dorazio RM (2009) Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Academic Press

Thomas L (1997) Retrospective power analysis. Conservation Biology 111: 276-280