Journal of Environmental Management 78 (2006) 194–199

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Why most conservation monitoring is, but need not be, a waste of time

Colin J. Legg[a,*], Laszlo Nagy[b]

[a]School of GeoSciences, University of Edinburgh, Crew Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JN, Scotland
[b]McConnell Ecological Research, 41 Eildon Street, Edinburgh EH3 5JX, Scotland

Received 1 July 2004; revised 16 March 2005; accepted 12 April 2005
Available online 19 August 2005

Abstract

Ecological conservation monitoring programmes abound at various organisational and spatial levels from species to ecosystem. Many of them suffer, however, from the lack of details of goal and hypothesis formulation, survey design, data quality and statistical power at the start. As a result, most programmes are likely to fail to reach the necessary standard of being capable of rejecting a false null hypothesis with reasonable power. Results from inadequate monitoring are misleading for their information quality and are dangerous because they create the illusion that something useful has been done. We propose that conservation agencies and those funding monitoring work should require the demonstration of adequate power at the outset of any new monitoring scheme.

Keywords: Conservation management; Monitoring; Power analysis; Statistical design

1. Objectives of monitoring

The aims of conservation management are either to maintain the *status quo* or to manipulate the system to achieve some predefined target by modifying the processes that are fundamental to ecosystem structure and functioning. Monitoring ['intermittent recording of the condition of a feature of interest to detect or measure compliance with a predetermined standard' (Hellawell, 1991)] is an essential tool in three main tasks: to inform the conservationist when the system is departing from the desired state; to measure the success of management actions; and to detect the effects of perturbations and disturbances.

2. Growth in monitoring as a conservation activity

Monitoring seems to be the automatic response of conservationists to any change or development that is seen as a potential threat to the environment, whether or not it is appropriate. A steep increase in the amount of monitoring work in the 1990s shows itself in the number of publications on the subject including, for example, the annotated bibliography on vegetation monitoring by Elzinga and Evenden (1997), which cites 1406 references. Many of the main conservation organisations are doing or commissioning monitoring work—but will the data that are being collected ever be of much use? There are undoubtedly good examples of long-term monitoring programmes collecting valuable data, but many projects seem unlikely to meet their stated objectives. Monitoring is often inadequate as, for example, Yoccoz et al. (2001), Byron et al. (2000), Wood et al. (2000) concluded in their reviews of the effectiveness of environmental impact statements and biodiversity monitoring. The results of inadequate monitoring can be both misleading and dangerous not only because of their inability to detect ecologically significant changes, but also because they create the illusion that something useful has been done (Peterman, 1990a). Such work may need to be repeated to a higher standard later with added costs.

Probable reasons for poor quality monitoring are not hard to find. One concerns the preferential use of qualitative or semi-quantitative monitoring techniques (e.g. recording of only presence/absence, estimation of population size/condition, site condition assessment), which may be adequate for some purposes, in place of quantitative methods (JNCC Common Standards Monitoring, http://www.jncc.gov.uk/page-2274; http://www.jncc.gov.uk/page-2282). The usual

* Corresponding author.
E-mail addresses: c.legg@ed.ac.uk (C.J. Legg), lnsc16324@blueyonder.co.uk (L. Nagy).

0301-4797/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2005.04.016

reason given for using qualitative methods is financial constraints. But the choice of a quantitative method does not guarantee success. A selection of textbooks on ecological monitoring and environmental impact assessment revealed that the majority of those examined (Southwood and Henderson, 2000; Glasson, 1999; Petts, 1999; Calow, 1998; Gilpin, 1995; Morris and Therivel, 1995; Wood, 1995; Goldsmith, 1991; Spellerberg, 1991; Fortlage, 1990) give little or no reference to the important issue of ensuring that the survey design is capable of detecting an impact on the system with adequate power. Only two of the books examined gave brief mention to the important question of hypothesis testing with references to more detailed methodology (Michener and Brunt, 2000; Treweek, 1999). Before the publication of the field manual on monitoring by Elzinga et al. (2001), more specialist books on research methodology (Ford, 2000; Krebs, 1989) or statistical analysis (Zar, 1999; Underwood, 1997; Cohen, 1988) had to be searched for an adequate treatment of the subject; this despite the fact that there are numerous excellent papers published in the ecological literature (see below). These papers seem to have been largely ignored by many practitioners. The reasons for this may be found in the inadequate coverage of monitoring design in degree and post-graduate courses, the lack of availability of suitable digests of scientific publications for in-service staff who commission monitoring work, and the lack of scientific peer-review of tenders by contract researchers (see for example Warnken and Buckley, 2000).

3. What are the requirements for a good monitoring programme?

It appears there is no cookbook recipe for the success and effectiveness of long-term studies. However, Strayer (1986) emphasised the importance of a simple and accommodating design in which the essential measurements and experimental treatments should be straightforward and unambiguously repeatable even by staff lacking sophisticated training (see Boxes 1 and 2).

Surveillance projects require good estimates of the accuracy and precision of parameters estimated. While this may also be true of monitoring, the ultimate-test of a

Box 1
Criteria for good management of a monitoring programme

- secure long-term funding and commitment
- develop flexible goals
- refine objectives
- pay adequate attention to information management
- train personnel and ensure commitment to careful data collection
- locations, objectives, methods and recording protocols should be detailed in the establishment report
- obtain peer review and statistical review of research proposals and publications
- obtain periodic research programme evaluation and adjust sampling frequency and methodology accordingly
- develop an extensive outreach programme

Based on Stohlgren (1995), Stewart et al. (1989), Hirst (1983).

Box 2
Recommendations for good design and field methods in monitoring

- take an experimental approach to sampling design
- select methods appropriate to the objectives and habitat type
- minimise physical impact to the site
- avoid bias in selection of long-term plot locations
- field markings must be adequate to guard against loss of plots
- ensure adequate spatial replication
- ensure adequate temporal replication
- blend theoretical and empirical models with the means (including experiments) to validate both
- synthesise retrospective, experimental and related studies
- integrate and synthesise with larger and smaller scale research, inventory, and monitoring programmes

Based on Yoccoz et al. (2001), Bakker et al. (1996), Stohlgren (1995), Stewart et al. (1989), and Strayer (1986).

good monitoring programme will collect data that provide sufficient information to reject the null hypothesis if it is false (see Box 3). Typical null hypotheses may be of one of the following forms:

1. 'the system has not changed beyond the predetermined limits of acceptable change'
2. 'the system has changed according to predetermined management objectives and is within the acceptable limits'

Box 3

Ways to increase power in monitoring

The power of a test depends on effect size, error variance, sample size and the Type I error rate (α). For example, the power of a t-test is derived from the t-distribution and the value of t given by:

$$t_{\beta(1),\upsilon} = \frac{\delta}{\sqrt{\frac{s^2}{n}}} - t_{\alpha,\upsilon}$$

The sample size required to detect a difference between means of δ with power $(1-\beta)$ is:

$$n = \frac{s^2}{\delta^2}(t_{\alpha,\upsilon} + t_{\beta(1),\upsilon})^2$$

(Note that t is a function of n so the solution must be obtained by iteration) where:

n is sample size,
s^2 is an estimate of variance,
δ is the minimum detectable difference,
$t_{\alpha,\nu}$ is the critical value of t for a probability of α (one-tailed or two-tailed as appropriate)
$t_{\beta,\nu}$ is the critical value of t for a one-tailed probability level β,
β is the probability of Type II error and
ν is the degrees of freedom.

Based on Zar (1999); Cohen (1988); Pearson and Hartley (1976), and Dixon and Massey (1969).

Effect size. The larger the effect, or the greater the change in the system, the easier the change will be to detect. Effect size can be increased by using more sensitive indicators, or by increasing the intensity of the treatment. However, when planning a monitoring programme the size of the effect is usually unknown. Power analysis, therefore, requires that the limits of acceptable change should be fixed at the planning stage and the monitoring designed so that a change of that magnitude will be detected if it occurs. There are no particular guidelines on how the limits of acceptable change should be fixed other than common sense (Toft and Shea, 1983), but see Cohen (1988).

Error variance. The power of a test depends on variability in the data. The greatest source of variability in the data in most cases stems from the fact that every sample unit is different from every other. This variance can be reduced at the design stage by, for example, increasing the size of the sample unit, stratification to reduce variance within strata, or the use of permanent plots, and by observer training (e.g. Pauli et al., 2004).

There is often an implicit assumption that different observers would obtain the same results when making observations. An estimate of the between-observer error is essential for long-term monitoring programmes where the same observer is unlikely to be responsible for the observations throughout the programme. In the few examples where between-observer and within-observer errors have been assessed for estimation of vegetation cover, for example, it has been found to be not insignificant (e.g. $\pm$10–20%, Nagy et al., 2002; Dethier et al., 1993; Kennedy and Addison, 1987; Sykes et al., 1983; Clymo, 1980).

Ecological systems may fluctuate from year to year because of chance events and changes in weather patterns and between-year variance cannot usually be assessed until the monitoring programme has been running for a few years. However, absence of this information is not grounds for ignoring power analysis at the design stage. There may be related studies available giving good estimates of the expected annual fluctuations and a power analysis based on an intelligent 'guesstimate' of between-year variance is considerably better than no power analysis at all.

Sample size. The simplest way to increase power is to increase sample size but this costs time and money and sample size should be traded-off against the quality of information that can be obtained from each observation. For example, estimates of plant cover made by averaging the visual estimates of cover in subunits within gridded quadrats show much less between-observer and within-observer error than visual estimates from ungridded quadrats. If the between-quadrat variance is high then large numbers of low-precision ungridded quadrats give greater power than the same amount of time spent on a few high-quality gridded quadrats (Legg, 2000; Nagy et al., 2002). Prior knowledge or a pilot study will be required to find the optimal method.

Manley (1992) suggested a practical approach to assessing required sample sizes. At first, a calculation is made about the maximum size of sample that can be collected given the resources available. From that, one can estimate the power of the test that one wishes to apply. If the estimated power is inadequate then one needs to decide whether to proceed or to abandon the study altogether as there is little point in a monitoring programme that cannot reject a null hypothesis that is false.

If large differences are to be detected the calculated sample size may be rather small. Statisticians caution that samples smaller than 20 may be too small to assume that the calculated variance would reasonably reflect population variance (Ebdon, 1985).

Type I error rate. By convention the Type I error rate (the probability of rejecting a true null hypothesis) is usually set arbitrarily at $\alpha = 0.05$, but increasing the acceptable Type I error rate can greatly increase the power of the test. This raises questions about the balance to be struck between Type I errors and Type II errors. For example, it has been proposed that the ratio of probability of Type I and Type II errors should equal the inverse of the ratio of the cost of the two errors (Di Stefano, 2001). In conservation ecology the cost of Type II errors—failure to reject the false null hypothesis—may be greater than the cost of Type I errors—rejection of a true null hypothesis (Shrader-Frechette and McCoy, 1992). Type II errors may mean the failure to detect damage to the resource and may result in loss of the resource. Type I errors mean that unnecessary additional management is applied—there is a cost implication, but the resource is not lost. A higher risk of Type I errors should, therefore, be accepted in order to increase the power of the test.

In the case of environmental threats where the costs of Type II errors are high the burden of proof should shift from the regulatory bodies to those causing the impact. However, the polluter must be required to demonstrate that the effect *does not* exceed acceptable limits *with high power* (Ebdon, 1985).

Type of test used. Monitoring programmes should be designed around a simple and powerful statistical model (e.g. analysis of variance, ANOVA) that can make use of all of the information available to reduce residual errors. Power can be increased by making assumptions about the data so that, for instance, parametric tests are usually more powerful than non-parametric tests. The hypotheses may also be refined; for example one-tailed tests are more powerful than two-tailed tests, although good *a priori* reasons must be present before one-tailed tests are used. Similarly, specifying planned means comparisons in ANOVA can increase power (Foster, 2001).

3. 'the perturbation of concern has had no impact on the system; all observed changes to the system can be attributed to other causes'

However, the third null hypothesis of 'no impact' will rarely be appropriate in ecology because it will almost always be false, even if the effect is exceedingly small (Johnson, 1999). What is interesting is not to know that the null hypothesis is false, but to ask if the change that has occurred is within acceptable limits. Null hypothesis 3 should, therefore, be re-written in most cases as 'the effects of the perturbation of concern do not exceed the limits of acceptable change'.

In all cases hypothesis testing requires not only accuracy and precision in the data but, most importantly, information about the statistical properties of the data; that is information about the *degree* of accuracy and precision.

4. Power analysis

The importance of estimating the power of a statistical test (power = 1.0 minus the probability of a Type II error, i.e. the probability of rejecting the null hypothesis when it is false) is well understood in the statistical literature (e.g. Zar, 1999; Sokal and Rohlf, 1995; Cohen, 1988). Nonetheless, the statistical power of ecological experiments is too rarely considered in the design stage (Nagy et al., 2002; Peterman, 1990a,b; Toft and Shea, 1983; Warnken and Buckley, 2000 and others). This may cause shortcomings in the interpretation of statistically non-significant results which are frequently (Peterman, 1990a), but erroneously, interpreted in the ecological literature as implying that the null hypothesis is true (i.e. that the perturbation has had no effect, the system is within the acceptable limits). But if the test has low power then there will be a high probability of a non-significant result even if the null hypothesis is false. Non-significant results may lead to the assumption that the perturbation has had no particular consequence when in fact there is serious loss of conservation value; inappropriate management may be continued even though the system has deviated well beyond the acceptable limits.

For the above reasons, power analysis is fundamental to the planning of long-term monitoring programmes because the consequences of inadequate design may not be obvious until the end of the programme by which time it will be too late to correct the problem. Whilst the importance of power analysis is being highlighted throughout this paper, the authors are mindful of the work of Hoenig and Heisey (2001) and of the warning by Fox (2001) about the need for '*due diligence, a mild degree of scepticism and appropriate attention to assumptions* [about distribution and error structure]' whilst performing power analysis.

It is all the more important that a power analysis is used to balance the risk of Type I and Type II errors against their respective costs in terms of both socio-economic and conservation objectives. For example, Di Stefano (2003) has recently argued that the frequently used rule of thumb of the 5 and 20% rates of Type I and Type II error are inappropriate. It is particularly important to select a design with high power when the cost of Type II errors is relatively high as has been pointed out by Field et al. (2004), quoting the example of monitoring panda populations.

The power of a test (Box 3) depends on several factors that are within the control of the observer: effect size (acceptable change); survey design and statistical test applied; sample size and the Type I error rate.

5. Statistical/process-model approach

Numerous authors emphasise the importance of developing a clear 'model' or hypothesis (e.g. Yoccoz et al., 2001; Humphrey et al., 1995; Pickett, 1991; Haug, 1983; Johnson and Bratton, 1978) so that the monitoring can be designed to test well-formed hypotheses using classical experimental approaches. 'These models, either explicit or subconscious, are part of every monitoring project and are usually characterised by being simple, correlated to causes, dynamic (incorporating temporal variability), discrete (reflecting periodic measurement), and analysed either statistically or by simulation' (Hirst, 1983).

A clear model has two important roles in the present context. Firstly, it focuses attention on the processes of change that are likely to be taking place. This will be important for identifying the best indicators that should be measured. The best indicators will be those that closely reflect the processes of change. Thus, measures of reproductive output or mortality may be more sensitive indicators than estimates of population size, provided that they remain relevant to the hypotheses. A common problem with conservation monitoring is the selection of suitable control sites and sites for valid replication as required by statistical analysis. The statistically balanced careful experimental design is rarely possible and the ecologist must make do with what is (or can be made) available. This means that one is often left comparing the system at the end of the monitoring programme with the baseline data, rather than comparing them with replicated control sites. Significant change in the system cannot then be automatically attributed to the impact of concern because all ecological systems change with time anyway. The change observed may be quite unconnected with the impact or management of interest. The frequently used BACI (Before–After-Control-Impact) design, although addressing some of these problems (Stewart-Oaten, 1992; Stewart-Oaten and Bence, 2001), has also been demonstrated as inadequate unless there is appropriate replication of sites (Underwood, 1994).

The second important role for a clear process-based model of the expected change is, therefore, to distinguish changes that are of no particular consequence, from changes that can be attributed to the impact or management treatment of interest. This can be achieved with an *a priori* model of the impact that makes precise predictions of the nature of the changes that should be expected. In the absence of a clear causal chain, a convincing case, therefore, requires that: results for several species follow a consistent pattern; plausible mechanisms for an ecological impact can be identified; and reasonable alternative mechanisms have been explored and ruled out (Schroeter et al., 1993). Several authors have emphasised the need for experimental work that must be conducted in association with the monitoring in order to provide and calibrate a model of the changes (e.g. Bakker et al., 1996; Strayer, 1986).

6. Conclusion

Few monitoring programmes pay sufficient attention to the details of hypothesis formulation, survey design, data quality and statistical power at the start. There is, therefore, a high probability that most monitoring will fail to reach the necessary standard of being capable of rejecting a false null hypothesis with reasonable power. It is the responsibility of the sponsoring bodies that commission the monitoring to ensure that a sufficiently high standard is maintained. 'When planning budgets managers should either give scientists sufficient funds and time to carry out a high power test of the null hypothesis, or not fund them at all' (Peterman, 1990a,b).

Acknowledgements

We thank Prof E.F. Bruenig, Dr A. Pullin, Dr N. Yoccoz and an anonymous referee for their comments on the manuscript.

References

Bakker, J.P., Olff, H., Willems, J.H., Zobel, M., 1996. Why do we need permanent plots in the study of long-term vegetation dynamics. Journal of Vegetation Science 7, 147–155.

Byron, H.J., Treweek, J.R., Sheate, W.R., Thompson, S., 2000. Road developments in the UK: an analysis of ecological assessment in environmental impact statements produced between 1993 and 1997. Journal of Environmental Planning and Management 43, 71–97.

Calow, P., 1998. Environmental Risk Assessment and Management. Blackwell Science, Oxford.

Clymo, R.S., 1980. Preliminary survey of the peat-bog Knowe Moss using varius numerical methods. Vegetatio 42, 129–148.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum, Hillsdale.

Dethier, M.N., Graham, E.S., Cohen, S., Tear, L.M., 1993. Visual versus random-point percent cover estimations: objective is not always better. Marine Ecology Progress Series 96, 93–100.

Di Stefano, J., 2001. Power analysis and sustainable forest management. Forest Ecology and Management 154, 141–153.

Di Stefano, J., 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology 17, 707–709.

Dixon, W.J., Massey, F.J., 1969. Introduction to Statistical Analysis. McGraw Hill, New York.

Ebdon, D., 1985. Statistics in Geography. Blackwell, Oxford.

Elzinga, C.L., Evenden, A.G., 1997. Vegetation Monitoring: An Annotated Bibliography, USDA, Forest Service, Intermountain Research Station INT-GTR-352.

Elzinga, C.L., Salzer, D.W., Willoughby, J.W., Gibbs, J.P., 2001. Monitoring Plant and Animal Populations. Blackwell, Oxford.

Field, S.A., Tyre, A.J., Jonzen, N., Rhodes, J.R., Possingham, H.P., 2004. Minimizing the cost of environmental management decisions by optimizing statistical thresholds. Ecology Letters 7, 669–675.

Ford, E.D., 2000. Scientific Method for Ecological Research. Cambridge University Press, Cambridge, Cambridge.

Fortlage, C.A., 1990. Environmental Assessment: A Practical Guide. Gower, Brookfield.

Foster, J.R., 2001. Statistical power in forest monitoring. Forest Ecology and Management 151, 211–222.

Fox, D.R., 2001. Environmental power analysis - a new perspective. Environmetrics 12, 437–449.

Gilpin, A., 1995. Environmental Impact Assessment (EIA): Cutting Edge for the Twenty-First Century. Cambridge University Press, Cambridge.

Glasson, J., 1999. Introduction to Environmental Impact Assessment: Principles and Procedures. UCL Press, London.

Goldsmith, F.B., 1991. Vegetation monitoring. In: Goldsmith, F.B. (Ed.), Monitoring for Conservation and Ecology. Chapman and Hall, London, pp. 77–86.

Haug, P.T., 1983. Resource inventory and monitoring under NEPA. In: Bell, J.F, Atterbury, T. (Eds.), Renewable Resource Inventories for Monitoring Changes and Trends: Proceedings of an International Conference. Oregon State University, College of Forestry, pp. 261–265.

Hellawell, J.M., 1991. Development of a rationale for monitoring. In: Goldsmith, F.B. (Ed.), Monitoring for Conservation and Ecology. Chapman and Hall, London, pp. 1–14.

Hirst, S.M., 1983. Ecological and institutional bases for long-term monitoring of fish and wildlife populations. In: Bell, J.F., Atterbury, T. (Eds.), Renewable Resource Inventories for Monitoring Changes and Trends: Proceedings of an International Conference. Oregon State University, College of Forestry, pp. 175–178.

Hoenig, J.M, Heisey, D.M., 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. The American Statistician 55, 19–24.

Humphrey, C.L., Faith, D.P., Dostine, P.L., 1995. Baseline requirements for assessment of mining impact using biological monitoring. Australian Journal of Ecology 20, 150–166.

JNCC Common Standards Monitoring http://www.jncc.gov.uk/page-2274;. http://www.jncc.gov.uk/page-2282.

Johnson, D.H., 1999. The insignificance of statistical significance testing. Journal of Wildlife Management 63, 763–772.

Johnson, W.C., Bratton, S.P., 1978. Biological monitoring in UNESCO biosphere reserves with special reference to the great Smoky Mountain National Park. Biological Conservation 13, 105–115.

Kennedy, K.A,, Addison, P.A., 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. Journal of Ecology 75, 151–157.

Krebs, C.J., 1989. Ecological Methodology. Harper Row, New York.

Legg, C.J., 2000. Review of Published Work in Relation to Monitoring of Trampling Impacts and Change in Montane Vegetation. Scottish Natural Heritage Review No. 131, Battleby.

Manley, B.F.J., 1992. The Design and Analysis of Research Studies. Cambridge University Press, Cambridge.

Michener, W.K., Brunt, J.W., 2000. Ecological Data: Design, Management and Processing. Blackwell, Oxford.

Morris, P., Therivel, R., 1995. Methods of Environmental Impact Assessment. UCL Press, London.

Nagy, L., Nagy, J., Legg, C.J., Sales, D.I., Horsfield, D., 2002. Monitoring vegetation change caused by trampling: a study from the Cairngorms, Scotland. Botanical Journal of Scotland 54, 191–207.

Pauli, H., Gottfried, M., Hohenwallner D., Reiter, K., Casale, R., & Grabherr, G., 2004. The GLORIA Field Manual - Multi-Summit Approach. European Commission, DG Research, EUR 21213, Official Publications of the European Communities.

Pearson, E.S., Hartley, H.O., 1976. Biometrika Tables for Statisticians, vol. 2. Charles Griffin, London.

Peterman, R.M., 1990a. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Science 47, 2–15.

Peterman, R.M., 1990b. The importance of reporting statistical power: the forest decline and acidic deposition example. Ecology 71, 2024–2027.

Petts, J., 1999. Handbook of Environmental Impact Assessment. Blackwell, Oxford.

Pickett, S.T.A., 1991. Long-term studies: past experience and recommendations for the future. In: Risser, P.G. (Ed.), Long-Term Ecological Research: An International Perspective. John Wiley and Sons, Chichester, pp. 71–85.

Schroeter, S.C., Dixon, J.D., Kastendiek, J., Smith, R.O., 1993. Detecting the ecological effects of environmental impacts-a case-study of kelp forest invertebrates. Ecological Applications 3, 331–350.

Shrader-Frechette, K.S., McCoy, E.D., 1992. Statistics, cost and rationality in ecological inference. Trends in Ecology and Evolution 7, 96–99.

Sokal, R.R., Rohlf, F.J., 1995. Biometry. Freeman, New York.

Southwood, T.R.E., Henderson, P.A., 2000. Ecological Methods. Blackwell, Oxford.

Spellerberg, I.F., 1991. Monitoring Ecological Change. Cambridge University Press, Cambridge.

Stewart, G.H., Johnson, P.N., Mark, A.F., 1989. Monitoring terrestrial vegetation for biological conservation. In: Craig, B. (Ed.), Proceedings of a Symposium on Environmental Monitoring in New Zealand with Emphasis on Protected Natural Areas. Department of Conservation, Wellington, pp. 199–208.

Stewart-Oaten, A., 1992. Assessing the effects of unreplicated perturbations: no simple solutions. Ecology 73, 1396–1404.

Stewart-Oaten, A., Bence, J.R., 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs 71, 305–339.

Stohlgren, T.J., 1995. Planning long-term vegetation studies at landscape scales. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, London, pp. 209–241.

Strayer, D., 1986. Long-Term Ecological Studies: An Illustrated Account of their Design, Operation, and Importance to Ecology. Occasional Publication of the Institute of Ecosystem Studies, New York Botanical Garden, No. 2.

Sykes, J.M., Horrill, A.D., Mountford, M.D., 1983. Use of visual cover estimates as quantitative estimators of some British woodland taxa. Journal of Ecology 71, 437–450.

Toft, C.A., Shea, .P.J., 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. American Naturalist 122, 618–625.

Treweek, J.R., 1999. Ecological Impact Assessment. Blackwell, Oxford.

Underwood, A.J., 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications 4, 3–15.

Underwood, A.J., 1997. Experiments in Ecology. Cambridge University Press, Cambridge.

Warnken, J., Buckley, R., 2000. Monitoring diffuse impacts: Australian tourism developments. Environmental Management 25, 453–461.

Wood, C., 1995. Environmental Impact Assessment. A Comparative Review. Longman, Harlow.

Wood, C., Dipper, B., Jones, C., 2000. Auditing the assessment of the environmental impacts of planning projects. Journal of Environmental Planning and Management 43, 23–47.

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

Zar, J.H., 1999. Biostatistical Analysis. Prentice Hall, Englewood Cliffs.