

# Monitoring of biological diversity in space and time

Nigel G. Yoccoz, James D. Nichols and Thierry Boulinier

**Monitoring programmes are being used increasingly to assess spatial and temporal trends of biological diversity, with an emphasis on evaluating the efficiency of management policies. Recent reviews of the existing programmes, with a focus on their design in particular, have highlighted the main weaknesses: the lack of well-articulated objectives and the neglect of different sources of error in the estimation of biological diversity. We review recent developments in methods and designs that aim to integrate sources of error to provide unbiased estimates of change in biological diversity and to suggest the potential causes.**

The term 'monitoring' has been used to describe many types of activities. Here, we define monitoring as the process of gathering information about some system STATE VARIABLE(S) (see Glossary) at different points in time for the purpose of assessing system state and drawing inferences about changes in state over time. Because we focus on the monitoring of biological diversity, the systems of interest are typically ecosystems or components of such systems (e.g. communities and populations), and the state variables of interest include quantities such as species richness, species diversity, biomass and population size.

In recent years, most developed countries have established monitoring programmes for biological diversity<sup>1–3</sup>. The scale, design and intensity of these programmes vary enormously. Although many programmes are restricted to a few subjectively selected sampling sites, in 1992, the US National Resources Inventory had almost one million sampling points chosen according to an elaborate stratified random-sampling scheme<sup>1</sup>. Obviously, the extent and strength of the inferences drawn will vary depending on the design used. However, many existing monitoring programmes suffer from various design deficiencies. Specifically, many programmes appear to be developed without paying adequate attention to three basic questions: (1) Why monitor? (2) What should be monitored? and (3) How should monitoring be carried out?

With respect to 'why monitor?', Krebs<sup>4</sup> wrote: 'Monitoring of populations is politically attractive but ecologically banal unless it is coupled with experimental work to understand the mechanisms behind system changes'. Monitoring is often explicitly linked to the assessment of management policies (with respect to implementation and expected effects); however, the role of research in framing actions and operational definitions of the management objectives, as well as in identifying the links between actions and effects to improve future

management, is often ignored. Researchers have therefore argued for a better integration of monitoring, management and research<sup>5</sup>. In particular, the use of adaptive management, where management and monitoring are explicitly used to gain reliable knowledge about the ecological system and to thus reduce uncertainty, represents a promising<sup>6,7</sup>, but challenging, perspective. With respect to the question of how to monitor, many existing programmes either ignore or deal ineffectively with two primary sources of variation in monitoring data, spatial variation and DETECTABILITY. A recent review of vertebrate monitoring programmes found substantial shortcomings in most of the programmes<sup>8</sup>.

Here, we focus on the why, what and how of monitoring programmes, and provide recommendations for the design and conduct of such programmes. We recognize that some of these recommendations are technical, whereas others involve general issues, but we believe them all to be important determinants of the ultimate utility of monitoring results, and thus worthy of serious consideration.

## Why monitor?

It is difficult to find explicit statements of objectives for many monitoring programmes. Indeed, the rationale underlying the development of many programmes seems to be the simple idea that additional information about any system will be inherently useful. It is much more efficient to specify clearly the objectives of a proposed monitoring programme. Such objectives can be usefully categorized into two general classes – scientific and management. Scientific objectives focus entirely on learning and developing an understanding of the behaviour and dynamics of the monitored system. Monitoring programmes designed to aid management provide information that is useful in making informed management decisions.

## Scientific objectives

Different approaches can be used to develop an understanding of system behaviour and dynamics (i.e. to meet scientific objectives) from monitoring data. The approach that yields the strongest inferences involves monitoring in conjunction with manipulation of the studied system for the specific purpose of testing or evaluating hypotheses of interest<sup>9</sup>. Competing *a priori* hypotheses are used to

### Nigel G. Yoccoz\*

Division of Arctic Ecology,  
Norwegian Institute for  
Nature Research, Polar  
Environmental Centre,  
N-9296 Tromsø, Norway.  
\*e-mail: nigel.yoccoz@  
ninasos.ninaniku.no

### James D. Nichols

US Geological Survey,  
Patuxent Wildlife  
Research Center, Laurel,  
MD 20708, USA.

### Thierry Boulinier

Laboratoire d'Écologie,  
Université Pierre et Marie  
Curie, CNRS UMR 7625,  
7 Quai Saint-Bernard,  
75252 Paris Cedex 05,  
France.

develop predictions about the changes in state variables that are expected to result from the manipulation(s). Inference is then based on comparing these predictions with estimates of these variables that come from monitoring data before and after the manipulation.

A more common approach to the use of monitoring data to serve scientific objectives is based on retrospective analyses; for example, it can rely on the statistical analyses of time series of population abundance or community-level state variables. Such analyses permit only weak inferences, in the sense that more than one hypothesis can explain the observed pattern<sup>10</sup>. For example, when a downward trend is recorded in a biological diversity component, it is easy to find positive trends in multiple variables that reflect environmental degradation (e.g. decreases in many natural habitats that result from land use changes, increases in a variety of environmental contaminants, or changes in numbers of predators or competitors). However, although such pairs of variables will often exhibit significant statistical correlations, such associations do not provide strong evidence of cause–effect relationships. Confounding or correlation between various causal factors does not allow the investigator to disentangle their relative effects, and the low level of sampled variation for some potential factors results in low precision, even if these factors prove important because of future increases in the range of variation (e.g. climatic change<sup>11</sup>).

Despite the stronger inferences resulting from experimentation, manipulation is not always possible, and sometimes inference has to be based on retrospective analyses that rely on natural variation<sup>12</sup>. In such cases, the key to using monitoring for science is that *a priori* hypotheses should be used to develop predictions with which monitoring data can be compared, rather than using monitoring data only for *a posteriori* hypothesis generation (i.e. looking at an observed time series and then developing a story about the process that generated it). This recommendation does not mean that hypothesis generation cannot be useful and accidental learning cannot occur, but simply places emphasis on *a priori* thought and planning as a means of maximizing efficiency in times of limited resources.

#### Management objectives

Monitoring programmes developed in conjunction with management typically serve two specific functions: (1) identifying the system state, and (2) providing information on the system response to management actions. The identification of the system state is of obvious importance to management. For example, if the size of a managed population is believed to be too small, then management actions should be directed at increasing numbers. The use of data from monitoring programmes to learn about system response to management actions is similar to

the use of monitoring data to distinguish between competing scientific hypotheses. Management actions are viewed as manipulations of interest, and system responses identified by monitoring are compared with *a priori* predictions of alternative hypotheses about system response<sup>13,14</sup>.

Such use of monitoring programmes with management presumes *a priori* specification of management objectives, and these objectives should, in turn, be formulated using unambiguous variables, such as density or demographic rates of focal or indicator species and groups<sup>5,15</sup>. When the aim is to also understand the processes behind the changes observed, such as in ecosystem management<sup>16</sup>, monitoring the rates of change of some system state variables, such as the rate of species loss or the proportion of new species, can be of particular value. Non-quantitative state variables, such as ecosystem health, should be avoided. Monitoring can also be used to test how appropriate the state variables selected to characterize the system are. For example, when defining management objectives in terms of changes in the densities of indicator species, the programme should incorporate tests to ensure that the selected species are indeed indicators of the processes and variables of interest<sup>15</sup>.

#### What to monitor

Decisions about which variables to monitor are determined largely by the objectives of the monitoring programmes; that is, by the answer to ‘Why monitor?’ Monitoring programmes directed at scientific objectives should focus on the state variables and associated rate parameters that are important to the *a priori* hypotheses (and their associated models) of system behaviour. Monitoring programmes designed to inform management should focus on the state and other variables that are included in the OBJECTIVE FUNCTION, as well as on variables that are needed to model the managed state variables adequately<sup>13</sup>. For example, in the harvest management of North American mallard ducks *Anas platyrhynchos*, population size of mallards is the state variable of interest in the objective function, but the monitoring programmes also includes mallard harvest (which is not a system state variable but which does appear in the objective function) and the number of wetlands on key breeding areas (which influences mallard abundance)<sup>6,17</sup>.

Biological diversity can be studied and managed at levels of biological organization ranging from genes to ecosystems. We restrict this article to species diversity, often referred to as ecological diversity. Ecologists have developed an almost endless number of diversity indices (e.g. Shannon-Weaver and Simpson; Box 1), but most of them can be seen as weighted sums of the relative abundances of species. Whereas original measures of species diversity focused primarily on the relative abundances of each species, recent proposed measures have incorporated

### Box 1. Measures of ecological diversity

Many measures of species, or ecological, diversity ( $D$ ) can be seen as special cases of a general weighted sum of the relative abundances of different species.

Therefore, they are closely related, and can be written generally as (Eqn I)<sup>a</sup>:

$$D = \sum_{j=1, \dots, k} R_j(\boldsymbol{\pi}) \pi_j \quad [I]$$

where  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$  and is the vector of relative abundances,  $k$  is the number of species and  $R_j(\boldsymbol{\pi})$  represents a rarity measure for the  $j$ -th species. Defining  $R_j(\boldsymbol{\pi})$  as (Eqn II):

$$R_j(\boldsymbol{\pi}) = (1 - \pi_j^\beta) / \beta \quad [II]$$

where  $\beta \geq -1$ , gives the classic measures of ecological diversity: species richness minus one (for  $\beta = -1$ ), Shannon-Weaver index (by taking the limit when  $\beta \rightarrow 0$ ) and Simpson index (for  $\beta = 1$ ). The choice of index should be made according to the goals of the monitoring programme, and classic measures might not be adequate. In particular, the weights  $R_j(\boldsymbol{\pi})$  could be

modified to take into account the economic values of the species, as well as their taxonomic values (e.g. endemic versus introduced species). An extreme case is when diversity is defined only with respect to a few flagship or umbrella species and when methods used to monitor populations instead of communities can be directly implemented<sup>b</sup>.

More general measures of biological diversity will take into account explicitly the distance between species [e.g. through the ordination of species traits<sup>c,d</sup> (functional or traits diversity) or measured along phylogenetic trees]. For example, a phylogenetic measure might be defined as (Eqn III)<sup>e</sup>:

$$D = [\sum_{i < j} \omega_{ij} \pi_i \pi_j] / [k(k-1)/2] \quad [III]$$

where  $\omega_{ij}$  is the phylogenetic path length between species  $i$  and  $j$  traced through a classification of the full set of species. Such an approach would give less importance to a species belonging to a

species-rich genus than to a species being the unique member of a family.

An estimate of  $D$  is obtained by using the estimates of  $\boldsymbol{\pi}$  in Eqn III. The main difficulties with this estimation stem from the fact that  $k$  is not known, that the contributions of unseen species to the diversity must, therefore, be estimated, and that the estimate of  $\boldsymbol{\pi}$  might be quite poor (low precision and/or large bias; Box 2).

#### References

- a Barabesi, L. and Fattorini, L. (1998) The use of replicated plot, line and point sampling for estimating species abundance and ecological diversity. *Environ. Ecol. Stat.* 5, 353–370
- b Thompson, W.L. *et al.* (1998) *Monitoring Vertebrate Populations*, Academic Press
- c Solow, A.R. and Polasky, S. (1994) Measuring biological diversity. *Environ. Ecol. Stat.* 1, 95–107
- d Dolédec, S. *et al.* (1999) Species traits for future biomonitoring across ecoregions: patterns along a human-impacted river. *Freshw. Biol.* 42, 737–758
- e Clarke, K.R. and Warwick, R.M. (1999) The taxonomic distinctness measure of biodiversity: weighting of step lengths between hierarchical levels. *Mar. Ecol. Progr. Ser.* 184, 21–29

other aspects, such as ecosystem values, economic values or taxonomic distinctness<sup>18,19</sup>. For example, rare or endangered species would probably be weighted more heavily than are common species in objective functions for conservation management. This position has been criticized on ethical grounds<sup>20</sup>, but, given that resources are limited and that not all species have the same role in the ecosystem or the same importance in terms of ECOLOGICAL SERVICES, such a weighting might be justified. Because management objectives are value dependent, these value-weighted diversity metrics are especially suitable for incorporation into formal objective functions. In addition, multiple measures to reflect different components of biological diversity should be monitored for some management purposes.

The sampling design for the monitoring programme will, in turn, be dependent on the choice of biological diversity measures. For some objectives, it might be adequate to focus on species richness for some group(s) as the state variable(s) of interest. However, for management objectives such as those ascribing different values to individuals of different species, it will be necessary to estimate the abundances of each species in the community of interest. Abundances of multiple species typically require more effort to estimate than does species richness, so decisions about what variables to monitor are important determinants of the design and expense of the programme<sup>21</sup>.

Both scientific and management programmes recognize the importance of dynamic processes, focusing on responses of state variables to environmental variation and, in the case of managed systems, to different management actions. Measures of biological diversity focused on system functioning, rather than solely on system state, are therefore increasingly preferred<sup>16</sup>. It is also important that monitoring programmes provide data to estimate not only the state variables of interest, but also the rate parameters that determine system dynamics. Boulinier *et al.*<sup>22</sup> studied the association between landscape fragmentation and rates of turnover, local extinction and colonization of avian communities. They showed that the lower species richness and higher temporal variability in species richness of forest birds observed in more fragmented landscapes could be explained by higher rates of local extinction and turnover (i.e. the proportion of new species).

#### How to monitor

There are two potential sources of error that should be considered when estimating biological diversity<sup>8,23,24</sup>.

#### Detection error

The first source of error occurs because few survey methods permit the detection of all individual animals, or even all species of animals, in surveyed areas. As stated previously, many methods of

**Box 2: The use of count statistics in monitoring: indices, detectability and estimation**

There are two approaches to the use of count statistics representing the number of individual animals or species counted in a population or community. The first involves the estimation of detectability, the other the use of count statistics as indices. Such count statistics are denoted as  $C_i$  (i.e. the number of individual animals or species counted at time or place  $i$ ). For the purposes of illustration, consider the problem of estimating species richness and think of  $C_i$  as denoting the species counted at time or place  $i$  (the same reasoning also extends to the problem of abundance estimation). The relationship between the quantity of interest ( $N_i$  = the number of species in the sampled community) and the actual count ( $C_i$ ) can be written as (Eqn I):

$$E(C_i) = N_i p_i \quad [I]$$

where  $E(C_i)$  denotes the expected value of the random variable  $C_i$ , and  $p_i$  is the probability that a member of  $N_i$  is detected and thus appears in the count statistic,  $C_i$ . This detection probability can also be viewed as the expected proportion of species detected at  $i$ .

If the detection probability associated with a count statistic can be estimated,

species richness can be estimated as follows (Eqn II):

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \quad [II]$$

where  $\hat{\phantom{x}}$  denotes estimates. Virtually all methods for estimating species richness and animal abundance follow the general form of Eqn II (Refs a,b).

The other approach uses count statistics directly for estimation. The count statistics are referred to as 'indices', and detection probabilities are assumed to be constant over time and space, a strong assumption<sup>c</sup>.  $\lambda_{ij}$  is defined as  $N_j/N_i$  and is the ratio of species richness at two different times or places,  $i$  and  $j$ . If  $i$  and  $j$  are two different points in time ( $i < j$ ), then  $\lambda_{ij}$  is often labelled 'trend', whereas if  $i$  and  $j$  indicate two points in space, then  $\lambda_{ij}$  is often referred to as relative richness. When count statistics are treated as indices, then  $\lambda_{ij}$  is typically estimated as the ratio of count statistics (Eqn III):

$$\lambda_{ij} = \frac{C_j}{C_i} \quad [III]$$

However, the expected value of this estimator can be approximated (using Eqn I) as (Eqn IV):

$$E(\lambda_{ij}) \approx \frac{E(C_j)}{E(C_i)} = \frac{N_j p_j}{N_i p_i} \quad [IV]$$

Examination of Eqn IV indicates that the estimator in Eqn III is a reasonable one (i.e. is approximately unbiased, but see Ref. d) only when  $p_i = p_j$ , that is, when the detection probabilities at times or places  $i$  and  $j$  are equal. If the detection probabilities themselves are viewed as random variables, then the assumption required for Eqn IV to be a reasonable estimator is:  $E(p_i) = E(p_j)$ .

**References**

- a Bunge, J. and Fitzpatrick, M. (1993) Estimating the number of species: a review. *J. Am. Stat. Assoc.* 88, 364–373
- b Lancia, R.A. *et al.* (1994) Estimating the number of animals in wildlife populations. In *Research and Management Techniques for Wildlife and Habitats* (Bookhout, T., ed.), pp. 215–253, The Wildlife Society
- c Conroy, M.J. and Nichols, J.D. (1996) Designing a survey to assess mammalian diversity. In *Measuring and Monitoring Biological Diversity. Standard Methods for Mammals* (Wilson, D.E. *et al.*, eds), pp. 41–49, Smithsonian Institution Press
- d Barker, R.J. and Sauer, J.R. (1992) Modelling population change from time series data. In *Wildlife 2001: Populations* (McCullough, D.R. and Barrett, R.H., eds), pp. 182–194, Elsevier Science

estimation of species diversity require measures of relative abundance. Ideally, estimates of local diversity would be based on unbiased estimates of the abundances of all species present locally (Box 1). Distance sampling and capture–recapture are two classes of method that can be used to estimate the detection probabilities associated with count statistics (animals observed and captured, respectively)<sup>8</sup>, and thus to estimate abundance (Box 2). However, such methods are often claimed to be impractical or too expensive to be implemented for all species. Therefore, most programmes rely on indices of population abundance (i.e. count statistics assumed to be related to true abundance), such as tracks or the number of individuals caught, heard or seen<sup>2,3</sup>. Translation of such indices into estimates of species diversity assumes that ratios of indices estimate relative abundance; that is, that individuals of all species are equally detectable and that all species are detected.

Species richness is the state variable of interest in some monitoring programmes, but the use of index statistics for this purpose again assumes that all species are detected or at least are detected with equal probability. Regardless of whether species

richness or species diversity metrics based on relative abundances are the state variables of interest, monitoring programmes should not simply rely on unadjusted count statistics or indices as a means of drawing inferences about temporal change. Instead, monitoring programmes should incorporate efforts to estimate the detection probabilities associated with the selected count statistics and survey methods<sup>8</sup>, even if such efforts are restricted to a subsample of surveyed locations. Standardization of sampling effort and observer effects might help in reducing sources of heterogeneity, but will not eliminate it, because other variables will influence species detectability. Without some knowledge of detection probabilities associated with collected count statistics (Box 2), it is not possible to draw strong inferences about the monitored system<sup>24–26</sup>.

**Spatial variation and survey error**

The second source of error in diversity estimation involves the inability to survey large areas entirely, and the resulting need to draw inferences about large areas based on samples of locations within those areas. Many issues related to management and biological diversity focus on spatial variation in

### Box 3: Probability sampling, judgement sampling and unbiased inferences

Relatively few monitoring programmes focus explicitly on biological diversity. However, many large-scale programmes have investigated changes and associated causes of natural resources in general. Olsen *et al.*<sup>a</sup> contrasted the designs of such programmes and, in particular, their use of probability-based sampling. Many programmes rely on judgement sampling, with samples chosen on the basis of their assumed representativeness of potential changes or sometimes pure practical convenience. Such sampling designs do not allow for unbiased estimation of trends, and, in a few cases where both probability and judgement sampling were used simultaneously, judgement sampling has been shown to result in substantially biased estimates<sup>b</sup>. Some of the biases might be corrected using model-based inferences for estimating changes in ecological diversity, instead of the design-based inferences used in classic sampling works<sup>c</sup>. Model-based inferences have advantages – providing explicit models for the spatial distribution of diversity as well as the factors affecting it – but also disadvantages, because the inferences might be sensitive to the model assumptions.

As an example, species distributions could be modelled as a function of spatial variables such as percentage cover of a given habitat or fragmentation level. However, getting unbiased estimates of the relationships between habitat variables and species abundance depends on using a proper model structure, and, in particular, on including all relevant variables. Such a model, if it is a good approximation, has direct relevance for management<sup>c</sup>. Design-based inferences are not sensitive to having specified the model correctly, but cannot easily incorporate relationships between habitat variables and abundance or species richness.

#### References

- a Olsen, A.R. *et al.* (1999) Statistical issues for monitoring ecological and natural resources in the United States. *Environ. Monit. Assess.* 54, 1–45
- b Peterson, S.A. *et al.* (1999) Sample representativeness: a must for reliable regional lake condition estimates. *Environ. Sci. Technol.* 33, 1559–1565
- c Buckland, S.T. *et al.* (2000) Wildlife population assessment: past developments and future directions. *Biometrics* 56, 1–12

biological diversity. BIODIVERSITY HOTSPOTS could be used to define conservation strategies and reserve locations<sup>27</sup>. Gap Analysis is an attempt to use maps and models that predict animal distribution to identify how currently protected area networks cover the different habitat types and regional biodiversity<sup>28</sup>. However, most estimates of biological diversity are not based on an appropriate spatial sampling scheme, and thus do not ensure unbiased estimates of biodiversity at larger spatial scales. Indeed, many monitoring programmes focus on a few subjectively selected sites (sometimes called ‘sentinel sites’), which, in general, cannot be used to draw inferences about diversity or trends of larger regions (Box 3)<sup>29,30</sup>.

#### Obtaining better estimates

The general principles of sampling<sup>24,25</sup> also apply to the monitoring of biodiversity. This holds true regardless of whether the monitoring programme is established for scientific or management objectives. Monitoring programme design should include specification of the TARGET POPULATION, which is often

defined by the area or habitat that is the subject of scientific inference or management actions. Remote sensing is increasingly used to list all the units belonging to the target population, and it might be wise to define a target population as being wider than the current one if changes in the environment are expected. For example, monitoring wetlands might include surveying areas that represent potential wetlands, and forest monitoring might include areas that are right above the tree line<sup>26</sup> in anticipation of future warming. The list of units belonging to the target population or sampling frame is then used to select sampling units in conjunction with the sampling design (Box 4).

Statisticians have developed increasingly sophisticated sampling designs and analyses with wide applicability (Box 4). The important recommendations are that sample locations for monitoring programmes should be selected to permit inference to the larger area of interest (i.e. the intent is to make statements about the entire area of interest, even though only a sample of locations in this area have been surveyed) and that sampling designs should be chosen with respect to their efficiency, where efficiency refers to the precision of resulting estimates.

Given the potential importance of both sources of error (detection and spatial), we strongly recommend that they both be incorporated in estimates of biological diversity resulting from monitoring programmes. For example, if abundance of individuals in a particular species is of interest, then detectability should be estimated and used to estimate abundance at each selected sample plot. In the case of simple random sampling, the abundance estimate for the entire area can be viewed as the product of the mean abundance per plot and the total number of potential plots in the area (Box 5). In the case of unequal PROBABILITY SAMPLING, total abundance is estimated as the sum of the abundance estimates for each selected plot divided by the associated selection probabilities (Box 5). The variances of these estimates include components associated with both spatial variation and detectability.

We are aware of very few monitoring programmes that follow this suggested approach and attempt to deal with both spatial variation and detectability<sup>8</sup>. The development of such programmes would be efficient, because they not only provide more reliable estimates of parameters but also lead to a more precise evaluation of ecological hypotheses pertaining to spatial dynamics<sup>31</sup>.

An example of a large-scale monitoring programme designed to deal with both detectability and spatial variation is the May Aerial Survey for North American Ducks<sup>17,32</sup>. This survey covers a large portion of northern USA and Canada and includes the important breeding grounds for many species of North American ducks. With respect to spatial variation, the survey consists of aerial transects arrayed in a

#### Box 4: Recent developments in sampling designs over space and time

The choice of a sampling design depends on its efficiency and whether it can be practically applied. Although probability sampling ensures unbiasedness, the precision of the estimates of changes or effects of management policies will depend heavily on spatial patterns.

Stratification<sup>a,b</sup> will result in more efficient estimates if the area of interest (the target population) can be divided into relatively homogeneous areas. It might also ensure a better spread of samples over the area. Sampling can be done at more than one scale<sup>b</sup>; for example, considering the relationship between habitat variables and biodiversity defined at the scale of microhabitat versus patch sites. The US National Resources Inventory is based on such a two-stage sampling design, with sampling at a small scale being spatially restricted to ensure a better spatial spread of the units within a large-scale unit<sup>c</sup>. Various designs are implemented to estimate rates of change (i.e. sampling over time), and they usually combine independent random samples and repeated measurements on the same samples.

The actual characteristics of sampling designs might be effectively tailored to the monitored taxa or accessibility of sampling areas. Some flagship and rare species, such as large birds of prey or carnivores, use the same breeding areas repeatedly and are therefore relatively easy to census. However, because some new nesting sites or dens might be used each year, a sampling design should allow for a probability-based sampling of both the known breeding sites and potential new breeding sites – a dual-frame approach to sampling<sup>d</sup>.

Finally, most classic sampling designs will not be efficient for rare species. Recent work has focused on adaptive sampling designs<sup>e</sup>, where the intensity of sampling is dependent on initial sampling results.

#### References

- a Thompson, S.K. (1992) *Sampling*, John Wiley & Sons
- b Gimaret-Carpentier, C. *et al.* (1998) Sampling strategies for the assessment of tree species diversity. *J. Veg. Sci.* 9, 161–172
- c Olsen, A.R. *et al.* (1999) Statistical issues for monitoring ecological and natural resources in the United States. *Environ. Monit. Assess.* 54, 1–45
- d Haines, D.E. and Pollock, K.H. (1998) Estimating the number of active and successful bald eagle nests: an application of the dual frame method. *Environ. Ecol. Stat.* 5, 245–256
- e Thompson, S.K. and Seber, G.A.F. (1996) *Adaptive Sampling*, John Wiley & Sons

STRATIFIED, systematic sampling design, with transect density being proportional to duck density. Detectability is estimated via a double-sampling approach in which segments of aerial transects are also searched by ground crews. Detectability is estimated for each aerial survey crew (pilot plus observer) as the birds counted from the air:birds counted on the ground ratio. Annual estimates of abundance and associated variances properly incorporate spatial variation (based on among-transect and among-strata variation in counts) and detectability (detectability estimates and their variances are used).

#### The future

Many monitoring programmes have been established with vague objectives that simply involve the collection of 'more information'. It is thus not surprising that many scientists consider monitoring biodiversity to be an unrewarding activity that involves little science. As we have suggested,

monitoring programmes can be established to meet scientific objectives, and such objectives are most likely to be met when the monitoring data are collected for the purpose of discriminating among competing *a priori* hypotheses about how the studied system works.

We believe that monitoring programmes are especially important to efforts to manage and conserve biodiversity. Given that monitoring programmes are frequently established to support management and that prediction of management consequences should be based on sound scientific knowledge of the system as well as on ecological principles and theories<sup>33</sup>, there needs to be a closer collaboration between scientists and managers. In particular, scientists should develop theories and field methods that fit with the scale and costs of management<sup>34</sup>.

An efficient framework is provided by the concept of adaptive management or policy design<sup>13</sup>. As with any effective management programme, adaptive management requires clear specification of both management objectives and management actions that are being considered. Adaptive management deals with scientific uncertainty by incorporating a set of models representing competing hypotheses about system responses to management. Monitoring data are used: (1) to assess the state of the system for the purpose of making periodic management decisions; and (2) to discriminate among models, or, in a Bayesian framework, to calculate changes in the relative probabilities or degrees of faith associated with the different models. Discrimination among competing models is based on comparing estimates of state variables from the monitoring programme with predictions of the competing models. In a Bayesian framework, model probabilities increase when model predictions correspond closely to estimates based on the monitoring programme. These probabilities decrease for models that predict poorly. Learning is reflected in changes in model probabilities over time and results in better management through more informed decisions. This approach has been successfully implemented for the management of the waterfowl harvest in North America<sup>6,17</sup>. A challenge for the future is to develop model sets that can be used for ecosystem management, such as for the interaction between large carnivores, ungulates and vegetation<sup>16</sup>.

#### Conclusion

We believe that many current monitoring programmes suffer from deficiencies associated with inadequate attention during programme design to the why, what and how of monitoring<sup>8</sup>. The general recommendation for those interested in establishing new monitoring programmes is that substantial thought should be devoted to the basic questions of 'how', 'what' and 'why'. Such a recommendation is also applicable for ongoing programmes, which, in some cases, could potentially be greatly improved by

### Box 5. Estimation of animal abundance: incorporation of both spatial variation and detectability

Some measures of biological diversity require estimates of abundance for the different species in the community. When abundance estimation is based on a monitoring programme covering a large area, it is important to incorporate two sources of error; that is detectability (Box 2) and spatial variation (Box 4). Consider a situation where a large area of interest is surveyed using simple random sampling<sup>a</sup>, and assume that there are  $K$  possible plots (sampling units) in the entire area of interest and that  $k$  of these are randomly selected with equal probability, given by  $P = k/K$ . An unbiased estimator of abundance for the entire area is given by Eqn I:

$$\hat{N}_T = \frac{K}{k} \sum_{i=1}^k \hat{N}_i \quad \text{[I]}$$

where  $\hat{\phantom{x}}$  denotes estimators,  $N_T$  is the total abundance over the entire area, and  $N_i$  is the abundance estimate for sample plot  $i$ . The variance of this estimator can be written as (Eqn II)<sup>b</sup>:

$$\hat{v}ar(\hat{N}_T) = K^2 \left[ \frac{\left(1 - \frac{k}{K}\right) \hat{S}_{N_i}^2}{k} + \frac{\hat{v}ar(\hat{N}_i | N_i)}{K} \right] \quad \text{[II]}$$

where  $\hat{S}_{N_i}^2$  is the estimated plot to plot variance in abundance (Eqn III)

$$\hat{S}_{N_i}^2 = \frac{\sum_{i=1}^k (\hat{N}_i - \hat{N})^2}{(k-1)} \quad \text{[III]}$$

and  $\hat{N}$  is the mean abundance per plot (Eqn IV)

$$\hat{N} = \frac{\sum_{i=1}^k \hat{N}_i}{k} \quad \text{[IV]}$$

and  $\hat{v}ar(N_i | N_i)$  is the average sampling variance associated with the abundance estimates (Eqn V)

$$\hat{v}ar(N_i | N_i) = \frac{\sum_{i=1}^k \hat{v}ar(\hat{N}_i | N_i)}{k} \quad \text{[V]}$$

The first of the additive terms in brackets in Eqn II characterizes the variation associated with the selection of  $k$  of the  $K$  total plots for survey. When the plots are of equal size, then this component reflects spatial variation in animal abundance over the area. The second term reflects the average measurement error associated with the estimation of detection probability, and thus abundance, based on the count statistics (Box 2).

Because the first component of Eqn II involves plot-to-plot variation in abundance, the variance might be inflated when plots are of unequal size. In such cases, it might be reasonable to consider sampling plots with a probability proportional to size. If plots are sampled without replacement (once a plot has been selected it has no chance of appearing in the sample a second time), then total abundance can be estimated using the Horwitz–Thompson estimator<sup>a</sup> as (Eqn VI):

$$\hat{N}_T = \sum_{i=1}^k \frac{\hat{N}_i}{P_i} \quad \text{[VI]}$$

where  $P_i$  is the probability of including plot  $i$  in the sample (this probability might

be based on plot size or other characteristics). Note that, when selection probabilities are the same for all plots, ( $P_i = P$ ), this estimator (Eqn VI) equals the estimator under simple random sampling (Eqn I). If the sample size is predetermined, then the abundance estimator in Eqn VI has variance given in Ref. b. However, in many ecological surveys, sampling costs are fixed and depend on the plot sizes, thus sample size,  $k$ , is a random variable. In this case, the variance of Eqn VI can be estimated<sup>b</sup> as (Eqn VII):

$$\hat{v}ar(\hat{N}_T) = \sum_{i=1}^k \frac{(1 - P_i) \hat{N}_i^2}{P_i^2} + 2 \sum_{i=1}^k \sum_{j>i}^k \frac{(P_{ij} - P_i P_j) \hat{N}_i \hat{N}_j}{P_{ij} P_i P_j} + \sum_{i=1}^k \frac{\hat{v}ar(\hat{N}_i | N_i)}{P_i} \quad \text{[VII]}$$

The values of  $P_i$  and  $P_{ij}$  (probability that plots  $i$  and  $j$  are both in the sample of  $k$  selected plots) must be computed on a case by case basis (see example in Ref. b). The above expressions are written in terms of  $\hat{N}_i$  and are fairly general. Thompson<sup>a</sup> provides estimators and associated variances in terms of detection probability estimates,  $p_i$ , and count statistics,  $C_i$ , for certain designs and scenarios (e.g. simple random sampling with estimated detection probability assumed to be the same for each sample plot).

#### References

- a Thompson, S.K. (1992) *Sampling*, John Wiley & Sons
- b Skalski, J.R. (1994) Estimating wildlife populations based on incomplete area surveys. *Wildl. Soc. Bull.* 22, 192–203

reviewing, at different stages of their implementation, whether enough attention is given to these questions. With respect to ‘why’, significant progress has been made in developing reasonable hypotheses and associated models for the crucial processes underlying some patterns of biodiversity changes<sup>35,36</sup>.

The main difficulties in integrating such hypotheses and models into monitoring programmes are because some scientists consider that understanding processes at some locations is enough for predictions at larger scales, and because some managers fail to realize that neglecting processes

underlying the patterns of change will hinder the improvement of management policies. A primary recommendation for the development of future monitoring programmes is simply that designers should strongly consider the scientific and/or management objectives of the monitoring programme, and the manner in which resulting data are to be used to achieve these objectives.

With respect to ‘what’, the quantities to be estimated should be determined by the *a priori* hypotheses to be addressed by the programme and by the relative values of different components of diversity as specified in management objectives. With

## Glossary

Some of the terms used are technical, and we provide some guidance to the meaning of them below:

**Biodiversity hotspots:** areas with higher diversity than average

**Detection error:** uncertainty in the estimation of abundance, species richness, and so on, at a given sampling plot, resulting from the inability to detect every individual or species on the plot.

**Design-based inference:** inference based only on the features of the sampling design (e.g. sampling probabilities and stratification), avoiding strong assumptions about processes affecting the state of the system.

**Detectability:** probability that a member of a population of interest is detected during sampling.

**Ecological services:** benefits human populations derive, directly or indirectly, from the functioning of ecological systems.

**Model-based inference:** inference based on a statistical model specifying numbers and characteristics (sex, size, etc.) of individuals, populations or species in the system of interest as a function of other characteristics of the sample plots (e.g. their spatial proximity).

**Objective function:** a mathematical statement of management objectives, such as minimizing the probability of extinction over a specified time horizon.

**Probability sampling:** a formal sampling scheme to give every element (individual, population, etc.) a known positive probability of selection.

**Spatial or survey error:** uncertainty in the estimation of abundance, species richness, and so on, for an area of interest as a result of spatial variation among sampling plots.

**State variable:** variable within the system of interest that is used to characterize the system status.

**Stratification:** a sampling scheme for which the target population is partitioned into groups or strata, and the sampling is performed separately within each stratum.

**Target population:** entire set of sampling units to which findings of the survey are to be extrapolated.

## Acknowledgements

NGY thanks the Biodiversity Programmes of the Norwegian Research Council and the Polar Environmental Centre for support. The authors would like to thank J.M. Gaillard, R.A. Ims, X. Lambin, A. Loison and J.R. Sauer for useful comments on this article.

respect to the 'how' of monitoring, substantial progress has been made over the past decade in development and refinement of methods for estimating measures of state and change in local biodiversity<sup>8,21–24,37,38</sup>. These methods incorporate estimation of detectability and provide efficient sampling designs to address questions at different

spatial scales. Our recommendation is simply that future designs should incorporate both spatial sampling and detectability to obtain unbiased estimates of the relevant state variables. This will allow, in turn, for a better understanding of the spatial component of changes in biological diversity and the underlying causes.

## References

- Olsen, A.R. *et al.* (1999) Statistical issues for monitoring ecological and natural resources in the United States. *Environ. Monitor. Assess.* 54, 1–45
- Lindén, H. (1999) Measuring and monitoring the diversity of wildlife with the aid of the Finnish triangle scheme. *Suomen Riista* 45, 80–88
- Gilbert, G. *et al.* (1998) *Bird Monitoring Methods*, Royal Society for the Protection of Birds
- Krebs, C.J. (1991) The experimental paradigm and long-term population studies. *Ibis* 133, 2–8
- Noss, R.F. (1999) Assessing and monitoring forest biodiversity: a suggested framework and indicators. *Forest Ecol. Manage.* 115, 135–146
- Johnson, F.A. *et al.* (1997) Uncertainty and the management of mallard harvests. *J. Wildl. Manage.* 61, 202–216
- Innes, J. *et al.* (1999) Successful recovery of North Island kokako *Callaeas cinerea wilsoni* populations, by adaptive management. *Biol. Conserv.* 87, 201–214
- Thompson, W.L. *et al.* (1998) *Monitoring Vertebrate Populations*, Academic Press
- Walters, C.J. and Green, R. (1997) Valuation of experimental management options for ecological systems. *J. Wildl. Manage.* 61, 987–1006
- Nichols, J.D. (1991) Science, population ecology, and the management of the American black duck. *J. Wildl. Manage.* 55, 790–799
- Groisman, P.Y. *et al.* (1999) Changes in the probability of heavy precipitation: Important indicators of climatic change. *Clim. Change* 42, 243–283
- Green, R.H. (1979) *Sampling Design and Statistical Methods for Environmental Biologists*. John Wiley & Sons
- Walters, C.J. (1986) *Adaptive Management of Renewable Resources*, MacMillan
- Hilborn, R. and Mangel, M. (1997) *The Ecological Detective*, Princeton University Press
- Simberloff, D. (1999) The role of science in the preservation of forest biodiversity. *Forest Ecol. Manage.* 115, 101–111
- Boyce, M.S. (1998) Ecological-process management and ungulates: Yellowstone's conservation paradigm. *Wildl. Soc. Bull.* 26, 391–398
- Nichols, J.D. *et al.* (1995) Managing North American waterfowl in the face of uncertainty. *Annu. Rev. Ecol. Syst.* 26, 177–199
- Solow, A.R. and Polasky, S. (1994) Measuring biological diversity. *Environ. Ecol. Stat.* 1, 95–107
- Clarke, K.R. and Warwick, R.M. (1999) The taxonomic distinctness measure of biodiversity: weighting of step lengths between hierarchical levels. *Mar. Ecol. Prog. Ser.* 184, 21–29
- Wilson, E.O. (1992) *The Diversity of Life*, Penguin Press
- Williams, B.K. *et al.* *Analysis and Management of Animal Populations*, Academic Press (in press)
- Boulinier, T. *et al.* (1998) Higher temporal variability of forest breeding bird communities in fragmented landscapes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 7497–7501
- Skalski, J.R., and Robson, D.S. (1992) *Techniques for Wildlife Investigations*, Academic Press
- Thompson, S.K. (1992) *Sampling*, John Wiley & Sons
- Thompson, S.K. and Seber, G.A.F. (1996) *Adaptive Sampling*, John Wiley & Sons
- Fuller, W.A. (1999) Environmental surveys over time. *J. Agr. Biol. Environ. Stat.* 4, 331–345
- van Jaarsveld, A.S. *et al.* (1998) Biodiversity assessment and conservation strategies. *Science* 279, 2106–2108
- Flather, C.H. *et al.* (1997) Identifying gaps in conservation networks: of indicators and uncertainty in geographic-based analyses. *Ecol. Appl.* 7, 531–542
- Edwards, D. (1998) Issues and themes for natural resources trend and change detection. *Ecol. Appl.* 8, 323–325
- Stoddard, J.L. *et al.* (1998) Can site-specific trends be extrapolated to a region? An acidification example for the northeast. *Ecol. Appl.* 8, 288–299
- Bjørnstad, O.N. *et al.* (1999) Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends Ecol. Evol.* 14, 427–432
- Smith, G.W. (1995) *A Critical Review of the Aerial and Ground Surveys of Breeding Waterfowl in North America*, US Dept of the Interior, National Biological Survey, Biological Science Report 5
- Sinclair, A.R.E. (1997) Carrying capacity and the overabundance of deer: a framework for management. In *The Science of Overabundance: Deer Ecology and Population Management* (McShea, W.J. *et al.*, eds), pp. 380–394, Smithsonian Institution Press
- Bunnell, F.L. and Huggard, D.J. (1999) Biodiversity across spatial and temporal scales: problems and opportunities. *Forest Ecol. Manage.* 115, 113–126
- Soulé, M.E. and Terborgh, J., eds (1999) *Continental Conservation. Scientific Foundations of Regional Reserve Networks*, Island Press
- Brown, J.H. (1995) *Macroecology*, University of Chicago Press
- Wilson, D. *et al.* (1996). Introduction. In *Measuring and Monitoring Biological Diversity: Standard Methods for Mammals* (Wilson, D. *et al.*, eds), pp. 1–7, Smithsonian Institution Press
- Schwarz, C.J. and Seber, G.A.F. (1999) Estimating animal abundance: review III. *Stat. Sci.* 14, 427–456