

## Sampling in ecology and evolution – bridging the gap between theory and practice

Cécile H. Albert, Nigel G. Yoccoz, Thomas C. Edwards Jr, Catherine H. Graham, Niklaus E. Zimmermann and Wilfried Thuiller

*C. H. Albert (cecile.albert@m4x.org) and W. Thuiller, Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Univ. Joseph Fourier, Bp 53, FR-38041 Grenoble Cedex 9, France. – N. G. Yoccoz, Dept of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics, Univ. of Tromsø, NO-9037 Tromsø, Norway. – T. C. Edwards, Jr, U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit and Wildland Resources, Utah State Univ., Logan, UT 84322-5290, USA. – C. H. Graham, Dept of Ecology and Evolution, Stony Brook Univ., Stony Brook, NY 11794, USA. – N. E. Zimmermann, Land Use Dynamics, Swiss Federal Research Inst. WSL, CH-8903 Birmensdorf, Switzerland.*

Sampling is a key issue for answering most ecological and evolutionary questions. The importance of developing a rigorous sampling design tailored to specific questions has already been discussed in the ecological and sampling literature and has provided useful tools and recommendations to sample and analyse ecological data. However, sampling issues are often difficult to overcome in ecological studies due to apparent inconsistencies between theory and practice, often leading to the implementation of simplified sampling designs that suffer from unknown biases. Moreover, we believe that classical sampling principles which are based on estimation of means and variances are insufficient to fully address many ecological questions that rely on estimating relationships between a response and a set of predictor variables over time and space. Our objective is thus to highlight the importance of selecting an appropriate sampling space and an appropriate sampling design. We also emphasize the importance of using prior knowledge of the study system to estimate models or complex parameters and thus better understand ecological patterns and processes generating these patterns. Using a semi-virtual simulation study as an illustration we reveal how the selection of the space (e.g. geographic, climatic), in which the sampling is designed, influences the patterns that can be ultimately detected. We also demonstrate the inefficiency of common sampling designs to reveal response curves between ecological variables and climatic gradients. Further, we show that response-surface methodology, which has rarely been used in ecology, is much more efficient than more traditional methods. Finally, we discuss the use of prior knowledge, simulation studies and model-based designs in defining appropriate sampling designs. We conclude by a call for development of methods to unbiasedly estimate nonlinear ecologically relevant parameters, in order to make inferences while fulfilling requirements of both sampling theory and field work logistics.

### Principles and definitions of sampling

Answers to ecological and evolutionary questions (e.g. in genetics, functional ecology, biogeography) often require measuring relevant biological variables (e.g. presence, growth rate, functional trait, concentration) on items (e.g. individuals, populations, species, ecosystems) of interest in the system under study. The set of all possible items from which generalized processes and patterns are to be derived is called the target population (Fig. 1), a statistical meaning of “population” that should not be confounded with the biological one. A statistical target population can be uncountable, or can contain only a small fraction of items actually accessible for measurements, which makes it generally impossible to completely measure (census) all items. It is then necessary to define what is called a sampling frame (Fig. 1), which is the finite set of all items that could be measured, and to employ different sampling strategies to

obtain items to measure from the frame. The sampling frame ideally coincides with the target population, but may differ for various reasons such as accessibility, incompleteness, or size (Särndal et al. 1992). From this defined sampling frame, one can draw a subset or sample of items – called sampling units – to be measured (Fig. 1). Sampling units should be distinct and easy to define; however, there are many situations where sampling units are not easily identified. This is the case when studying species that do not have clear individuals (e.g. plants with vegetative growth) or when units are defined in terms of location or area (e.g. grid-cells of a map). In the latter case, there is an undetermined number of units that may or may not overlap.

Procedures to draw samples from the sampling frame, called sampling designs (Fig. 1), are of two main types: probabilistic vs non-probabilistic. These types differ in terms of estimation methodology and reliability (Lemeshow

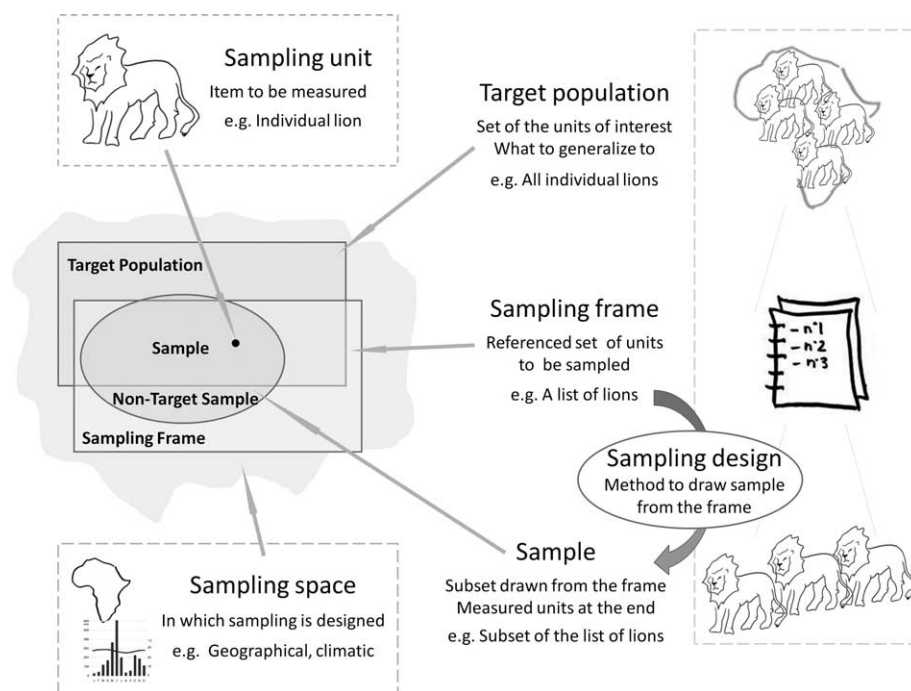


Figure 1. Sampling concepts and vocabulary.

and Levy 1999). To be reliable, a sampling design should lead to parameter estimates that are both unbiased (i.e. mean value of estimates equal to the target population value) and precise (i.e. small variability). This often requires a trade-off between bias and precision reached when the mean squared error ( $MSE = Bias^2 + Variance$ , Hastie et al. 2009) is minimized. In a probabilistic design, every sampling unit from the frame has a known non-zero probability of selection, leading to unbiased estimates of the mean and variance (and their confidence intervals) for the variables of interest. The unbiased Horvitz-Thompson estimator of the mean for the target population is then:

$$\sum_i \frac{1}{p_i} \times x_i \quad (1)$$

where  $x_i$  is the variable  $x$  measured on the  $i$ th sampling unit and  $p_i$  is the non-null probability to sample the  $i$ th sampling unit. In case of non-equal probabilities, measurements on units that are less likely to be sampled are given a higher weight, as they are expected to be under-represented in the observed sample. In a non-probabilistic design, some sampling units from the frame have no chance of being selected or their selection probability cannot be accurately determined. This can lead to biased estimates and may result in erroneous inferences (Lemeshow and Levy 1999).

Choice of a sampling method – which requires a proper identification of the target population, sampling frame, sampling units, sampling design, study extent and sampling effort – must be done within the frame of sampling theory. These choices influence the ecological or evolutionary patterns and processes that are revealed during analysis, the confidence in the resulting estimates, and the validity of extrapolation to other systems or time periods (Lemeshow and Levy 1999, Olsen et al. 1999, Skalski et al. 2005).

### Sampling in ecology

Sampling issues have always been an important aspect of ecological studies and are frequently discussed in ecological literature. For instance, different sampling designs can result in radically different species distribution models (Austin 1987, Hirzel and Guisan 2002, Austin et al. 2006), models for conservation prioritization (e.g. lichens, Edwards et al. 2004; butterflies, Haddad et al. 2008), economics forecasting (e.g. forests, Schreuder et al. 1999; fishes, Courbois et al. 2008, Hughes and Peck 2008), or public health initiatives (e.g. radioactivity monitoring, Scott et al. 2008). Though crucial, sampling issues are often difficult to overcome in ecological studies due to apparent inconsistencies between theory and practice, such as proper identification of target population and sampling units, or logistics problems (e.g. time, money, accessibility) when planning field work, or when working with information from many different extant datasets.

Unfortunately, this often leads to the implementation of simplified or so-called convenience sampling designs (Schreuder et al. 1999, Rosenstock et al. 2002), such as following transects or roads (e.g. US Breeding Bird survey, Peterjohn 1994, Bart et al. 1995), which, like all non-probabilistic designs, suffer from unknown biases. Yet even though probabilistic designs ensure the unbiased estimation of simple parameters like means and variances, they may be insufficient to fully address many ecological questions, particularly those estimating the strength of relationships between a response and a set of predictor variables over time and space. Extending the use of classical sampling principles on the estimation of means and variances to the evaluation of such relationships is not straightforward. Two key pitfalls concerning the appropriateness of sampling for evaluating these relationships include:

Truncated gradient: accurately estimating the shape of a biological pattern – such as the response of a functional trait to an environmental gradient, which in many cases will be non-linear – requires sampling across the entire range of variation of both the environmental and biological variables of interest (Mohler 1983, Austin 1987, Thuiller et al. 2010), and doing so in a balanced manner (i.e. not with a large amount of points in one part of the curve). Truncated gradients can lead to erroneous statistical inference, which in turn can affect ecological interpretation and lead to spurious predictions if the response is used to extrapolate a species distribution in space or time (Van Horne 2002, Thuiller et al. 2004). This may occur, for example, when using administrative instead of natural boundaries for delineating the study extent, or when samples do not exist over part of the range. It is then critical to employ a function that is constrained at its extremes and that cannot be extrapolated beyond these extremes. This common pitfall results when samples are drawn from a frame that misses a substantial portion of the target population.

Indirect gradients: inference from a given gradient assumes that a specific value of the gradient has the same meaning everywhere (Austin and Smith 1989). This assumption is unlikely in studies dealing with sites that are spread at continental or global scales, as rough global gradients (e.g. temperature) are less informative at local scales given they can subsume many underlying factors that are locally or physiologically important (e.g. soil, pollution, drought). This is also an issue for so-called indirect gradients (*sensu* Austin 1987), such as latitudinal and altitudinal gradients: an elevation of 1000 m a.s.l. does not have the same meaning in terms of environmental conditions in the Rocky Mountains, North America, and on Mount Kenya, Africa. Indirect gradients are inherently hard to interpret and are usually not transferable in time and space; however, they are extensively used in ecological studies in spite of explicit warnings concerning their use as surrogates of climatic gradients (Körner 2007, McCain 2009). Here, the sample frame may better coincide with the target population but be at a coarse grain that misses important, finer-grained effects.

## Objectives

Pitfalls concerning the appropriateness of sampling for explanatory or predictive model estimation can be avoided in many cases by: 1) an appropriate definition of the sampling space (e.g. geographical, climatic); 2) selection of a proper sampling design; and 3) the use of prior knowledge on the study system. The increased desire to obtain and use data over large, often continent-wide, spatial extents, coupled with ever-increasing availability of large electronic databases, makes it urgent to address these issues given how they affect the reliability of estimates and inferences made from field data and databases. We illustrate these three issues and provide suggested solutions obtained from a semi-virtual simulation study.

## Methods

### Simulation study

We used a semi-virtual simulation study to illustrate the consequences of sampling decisions on estimation of biological parameters. Semi-virtual studies are useful because they contain, and can hence depict, the “true” pattern of interest, and therefore can be used to test and compare sampling designs (Hirzel and Guisan 2002, Austin et al. 2006, Zurell et al. 2010). Our simulation study does not cover all sampling design possibilities, but covers five commonly implemented designs in ecological studies (see below). The simulation study was built in R version 2.10.1 (R Development Core Team 2008) using the packages *ade4*, *graphics*, *lattice*, and *VGAM*.

### Definition of the case study

For our semi-virtual simulation study we selected a study area with strong climate gradients that were not correlated with topography. This choice of a real landscape allows us to maintain realism in our simulations. The study area encompasses 30 000 km<sup>2</sup> located in southeastern France. It includes part of the French Alps and the Rhône Valley (Fig. 2). Sampling units were defined as pixels of the map (at 1 × 1 km) identified by their geographic coordinates (latitude and longitude). The area was further described by elevation data derived from the shuttle radar topography mission (SRTM3, <[www2.jpl.nasa.gov/srtm/](http://www2.jpl.nasa.gov/srtm/)>) and climate. For simplicity, we selected two uncorrelated climatic gradients from the French National Climatic Network (Aurhely, Benichou and Le Breton 1987): mean annual temperature and total annual precipitation. We defined a virtual climatically-constrained biological variable that varied between 0 and 100 and could be measured over the study area. This variable can be considered to represent one of many different ecological variable types, such as growth rate, habitat suitability, abundance, survival, radioactivity concentration, net primary productivity, or allele frequency (Fig. 2).

Because many biological variables show an optimum at some intermediate environmental condition and decline gradually towards the boundaries (Hirzel and Guisan 2002, Gaston 2003), the formulated variable followed a skewed Gaussian shape along both selected climatic gradients (obtained with the R function *dsnorm()*). To add some realism, and because two gradients may not be sufficient to describe the environmental conditions that influence the distribution of a biological variable, we added a random noise variable having a normal distribution (mean = 0, standard deviation = 4, i.e. normally spread between ca. –15 and 15) to simulate the other factors influencing the organism of interest (after Mohler 1983, Hirzel and Guisan 2002). We ensured that the optimum of the biological variable (highest value) occurred in the study area, but was not centered climatically (8.28°C and 1236 mm). To avoid the problems of sampling-induced bias due to the detectability of sampling units – an important topic not evaluated in our simulation – we assumed that the variable of interest had a detectability of 1.0, occurred everywhere, and is stable over time (Wintle et al. 2005, Mackenzie 2006).

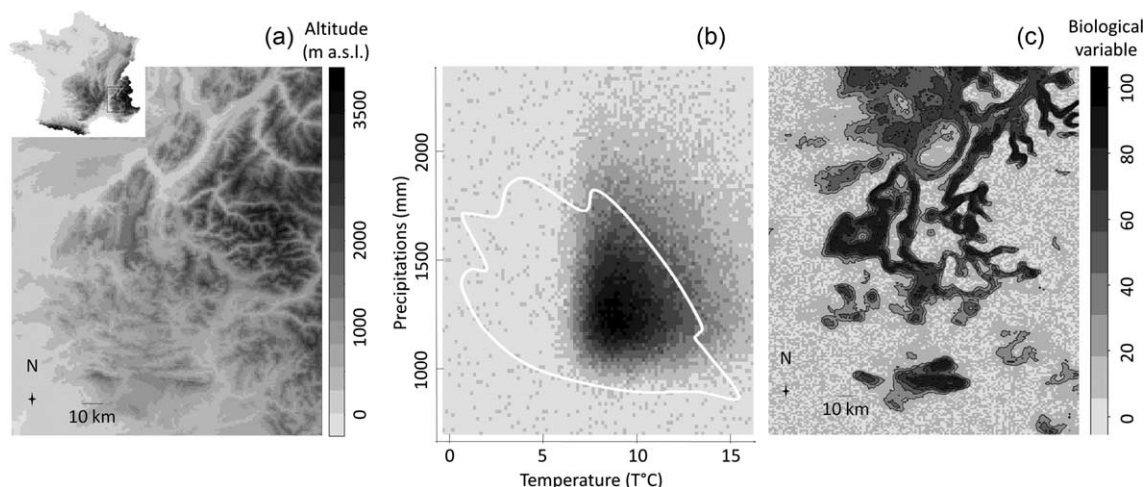


Figure 2. Study site and the virtual variable of interest. The study site is located in the south-east of France. (a) Altitudinal map of the study site including a part of the French Alps and of the Rhône valley. (b) Construction of the virtual biological variable as a skewed Gaussian curve along both temperature and precipitations gradients plus a noise. The white line delimits the environmental conditions existing in the study area. (c) Map of the virtual ecological variable through the study area. The virtual variable ranges from 0 to 100 (e.g. habitat suitability, relative biomass, relative fecundity).

The sampling frame is thus defined as the list of all the pixels in the defined geographic space.

### Virtual samplings

We tested the effects of five common sampling approaches implemented in ecological studies on our virtual biological variable. Samples were drawn from the sampling frame following designs of increasing complexity (see Fig. 3 for a visual example): 1) simple random, the simplest probabilistic design where each pixel from the frame had the same probability of being sampled. 2) Road-based sampling, a commonly used non-probabilistic design where only pixels close to the main roads had a non-null probability (i.e. inversely proportional to the distance to roads) of being sampled. 3) Stratified random by altitude, a commonly used probabilistic design where equal numbers of pixels were randomly drawn from a fixed number (here, ten) of altitudinal classes. 4) Systematic from the climatic space, a probabilistic design given a random start. Climatic space was stratified within a regular grid and one pixel was randomly drawn from each cell of this grid. 5) Model-based design, a learning non-probabilistic sampling process. We first drew a sub-sample from the frame using uniform sample from the climatic space (i.e. design 4). The response to the two climatic gradients was first estimated using a polynomial model with linear and quadratic terms for temperature and precipitation and an interaction on a first part of the sample:

$$\text{Variable} \approx \text{temperature} + \text{temperature}^2 + \text{precipitation} + \text{precipitation}^2 + \text{temperature} \times \text{precipitation} \quad (2)$$

Negative predictions were converted to zeros. This prior, model-based knowledge of how the response generally varies along environmental gradients was used to define the final sampling (second part), which followed a central composite factorial design centred on the estimated optimal conditions (OC) and surface response methodology (Box

and Hunter 1957, Wu and Hamada 2000). The two iterative parts of the overall sampling design were chosen here to contain an equal number of sampling units, but the optimal allocation is likely to depend on the problem.

To evaluate the effect of sample size relative to each design, samples containing 50, 200, or 1000 sampling units were drawn following each sampling design. Each combination of design and sample size was simulated 1000 times. For the three probabilistic designs (1, 3, 4), the probability  $p_i$  to sample the unit  $i$  was determined analytically. We ran the two non-probabilistic designs (2, 5) 10 000 times to numerically determine the probability  $p_i$  of the unit  $i$  of being sampled.

### Evaluation methods of the different samplings

We estimated two biological parameters from the virtual samples that we then compared to the “true” values: mean and optimum of the biological variable (Fig. 4). These are only two examples among all parameters that could be estimated for the variable (e.g. sum, variance, range, response to gradients). 1) We estimated the weighted mean (eq. 1, Horvitz-Thompson estimator) for the probabilistic samples, and without weights for the non-probabilistic samples. 2) We also estimated the optimal conditions (optimum) in environmental space for the biological variable from the linear model described above (Fig. 4c). Finally, we calculated the Euclidean distance (in standard deviation) between the “true” and the estimated optimum in the standardized variable space (e.g. the distance between the white circle and the white dots or black squares in Fig. 4c); this distance gives an idea of how well the optimum was estimated (Fig. 4b). We thus assumed that the variable could be estimated using a polynomial model (as in the eq. 2) with linear and quadratic terms for temperature and precipitations and an interaction.

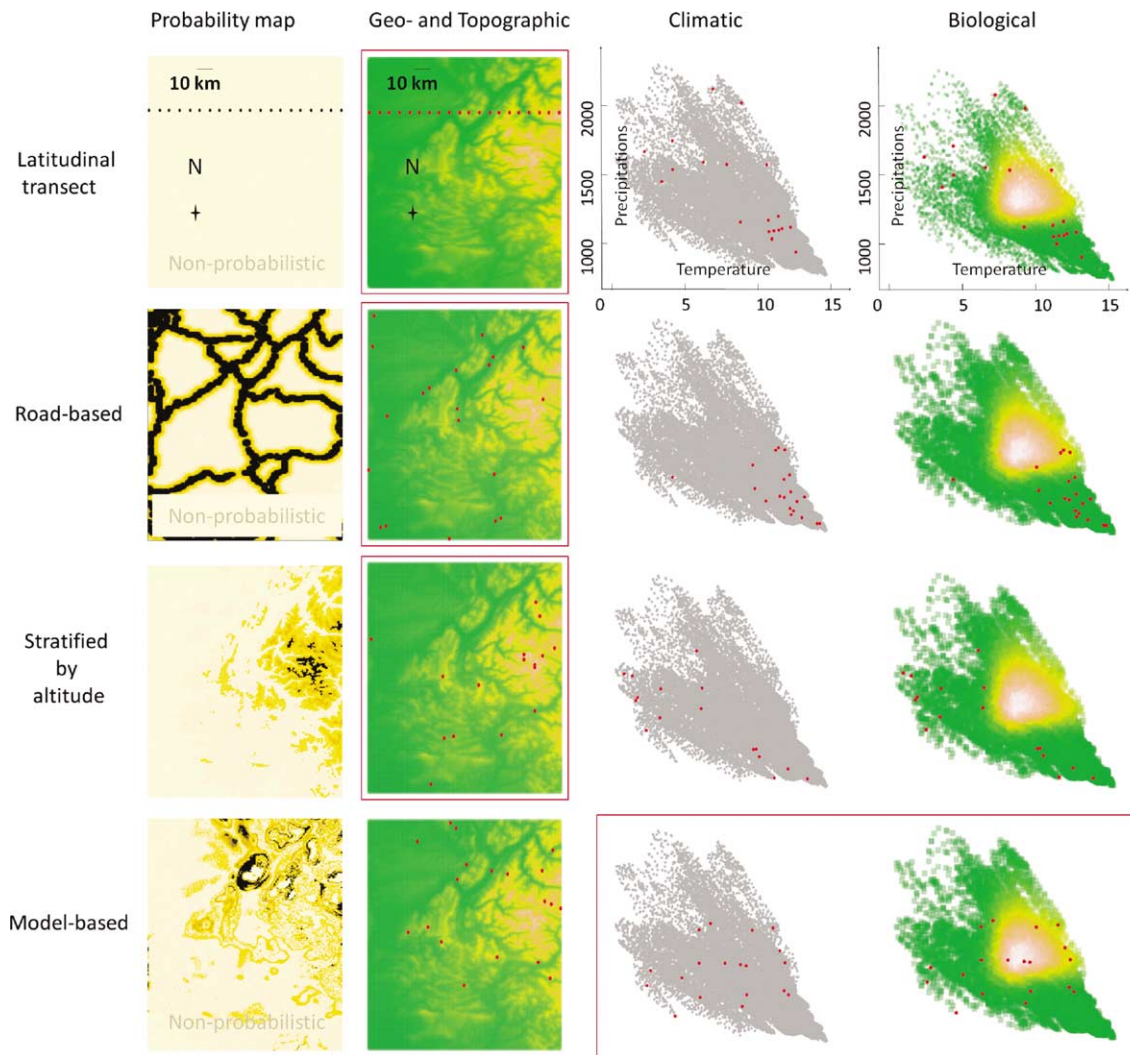


Figure 3. Importance of the sampling space onto which the sampling design is set up. Sampling designs used in this study: the red dots (20 measurements) represent a sample drawn from the frame through different sampling designs. From top to bottom, a latitudinal transect sampling (drawn from the frame within the geographic space), a road-based sampling (drawn from the frame within the geographic space), a stratification by altitude (drawn from the frame within the topographic space), a model-based sampling (drawn from the frame within the climatic space). The red boxes indicate in which space the sampling has been performed. Axes have not been repeated on all figures, they are the same within columns. Sampling probability (left column): patterns resulting from four different sampling designs, i.e. for each unit of the frame (here, each pixel from the map) its probability to be sampled for four contrasting sampling designs (the darker points represent the higher probabilities). These maps of probability are differing and will thus result in different weighting in follow-up calculations. Different spaces (from left to right): the topographic space represented within the geographical map (green for low altitudes and brown for high altitudes), the climatic space represented by temperatures and precipitations and the biological space represented by the variable of interest mapped into the climatic space (green for low values and brown for high values).

## Results

For both biological parameters, bias and precision of the estimates depend on the sampling design (Fig. 4). As expected, sample size had a strong effect on the precision of estimates of both the mean and the distance to optimum. In general, larger samples resulted in more precise estimates (Fig. 4). However, the efficiency of the sampling design also depends on the question and the parameters to be estimated (e.g. mean vs optimum, Fig. 4). The estimate of the mean was closer to truth (no bias and high precision) using the probabilistic samples where data were drawn using a simple random (1) and systematic sample from the climatic space

(4) than from any other sampling design (Fig. 4a). Unbiased estimates with lower precision were obtained with the stratification by altitude (3). In contrast, estimates were biased for the non-probabilistic designs. The mean was underestimated with the road-based design (2) because cold areas (higher values) were less frequently sampled (see truncated gradients), and was over-estimated with the model-based design (5), as sampling was more intense closer to the optimum.

However, the estimate of the variable's optimum was closer to the truth (smaller bias and higher precision) when calculated from the model-based design (5) than from any other sampling design (Fig. 4b). Two designs, systematic

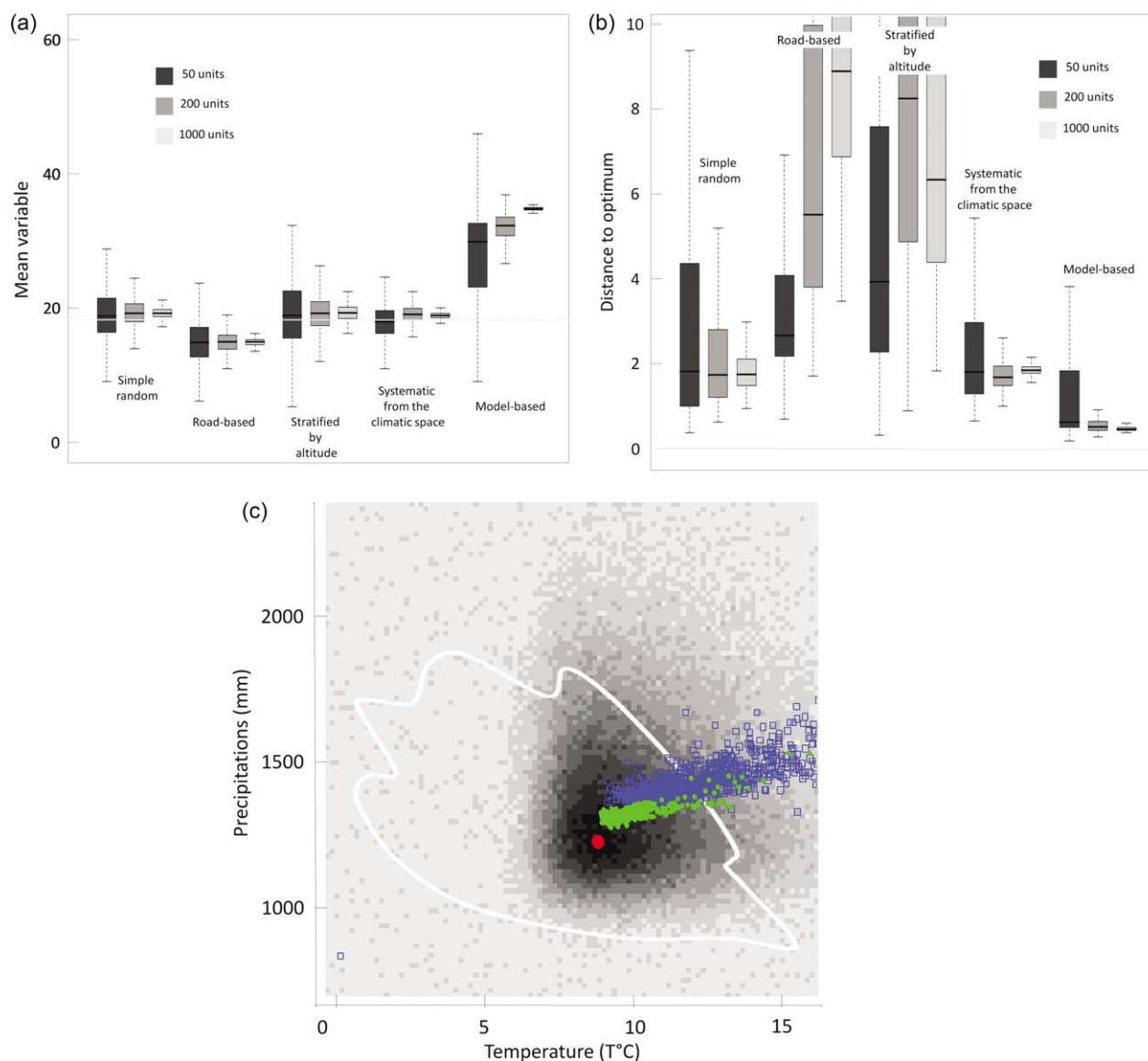


Figure 4. Effects of sampling designs and sample size on the precision of parameter estimations. Illustration of the sampling design effect on the precision and bias of the estimation of two biological variables: (a) the mean variable and (b) distance between the “true” optimum corresponding to environmental optimal conditions for the observed variable and the estimated one. Results are given for the five sampling designs: 1) simple random, 2) road-based, 3) stratified random by altitude, 4) systematic from the climatic space, and 5) a model-based. The boxplots of the estimated values give the median, and the first and third quartiles of the estimates (mean and distance to optimum). Results are given for 50, 200 and 1000 sampling units. (a) Distribution of the estimations of the parameter mean. The grey line is the true value on the studied area. (b) Distribution of the Euclidian distances between the true optimum of the parameter and the one estimated from virtual measurements. The grey line corresponds to 0, the best value that can be obtained. (c) Map of the true (red dot) and estimated (blue squares and green dots) optimum of the parameter, for 200 measures estimated optima for the simple random design (1) in blue and the model-based design (5) in green. The gray shade represents the virtual biological variable used in this study (Fig. 2b).

from the climatic space (4) and simple random (1), resulted in somewhat less precise estimates of the optimum. Finally, the road-based (2) and stratified random by altitude designs (3), resulted in very poor estimates of the response shape (Fig. 5) with its optimum (see truncated gradients). Interestingly, the higher the sampling effort, the poorer the estimates of the optimum, with mean estimates generally being greater than the target population value. With 200 measurements, the optimum was found at 19°C and 1758 mm for the simple random design compared to 9.8°C and 1376 mm for the

model-based design (Fig. 4c, 5). These were relative to the true optimums for temperature and precipitation of 8.28°C and 1236 mm, respectively.

## Discussion

### A story of sampling spaces

Our simulation study illustrates the importance of the selection of the appropriate sampling space (Fig. 3). The



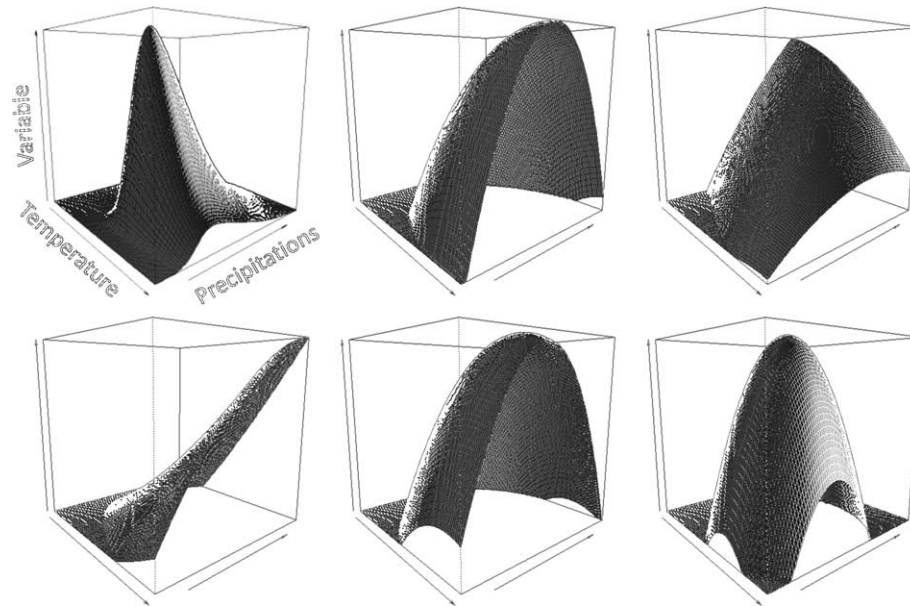


Figure 5. Estimation of the surface responses for the different designs. The 3D surface is given along the two environmental gradients temperature and precipitations. From top left to bottom right, the true surface and response curves obtained with 200 measurements for the simple random, road-based, the stratified random by altitude, the systematic from the climatic space, and the model-based designs.

sampling frame, here defined as the set of all geographical pixels in the study site, can be represented in different spaces (Hutchinson 1957): geographic space (e.g. latitude, longitude), topographic space (e.g. elevation), climatic space (e.g. temperature, precipitations), and biological space (e.g. habitat suitability). In our example, the sampling frame was defined in the geographic space. From this frame, we generated a set of sampling units that are not redundant in this space because one geographical location corresponds to only one possible unit. However, when projected into the climatic space, the frame looks less representative than in the geographic space, with a large percentage of the sampling units representing hot and dry conditions, and a small percentage of them representing other climatic conditions. Clearly, switching from one space to another can have consequences on the sampling (Fig. 3). For instance, the road-based sampling in our example leads to a complete coverage of the latitudinal gradient, but to an incomplete coverage of the altitudinal and temperature gradients, with a complete omission of cold mountainous areas far from roads. This easily explains why sampling in an inadequate space might result in problems as we described in the case of truncated gradients. It also shows that selecting the appropriate space to sample in is crucial for learning about the abiotic factors generally implied in ecological processes (Magnani 2009). The sampling is however often done in one space (e.g. geographical) while the pattern of interest occurs in a different space (e.g. climatic, Fig. 3). For example, both dispersal limitation and spatial autocorrelation of the environment (e.g. climate) can shape species distributions in the geographical space (Austin and Smith 1989, Araujo et al. 2009). Consequently, the selection of the sampling space may seem trivial on the surface but is not always evident in practice.

### Influence of the sampling design

#### *On the probabilities to sample a sampling unit*

How the sampling design may influence the patterns revealed from the sampled data is rarely discussed in ecological studies, where the focus is generally on the estimation of simple parameters such as population averages, size, or measures of diversity (Yoccoz et al. 2001, Edwards et al. 2006). However, the method used to implement the sampling can strongly influence the estimates by affecting the selection probabilities of the sampling units (Fig. 3, Austin and Smith 1989, Sauer et al. 2003). Sampling probabilities are determined analytically in the case of probabilistic designs, and evaluated numerically in the case of non-probabilistic designs. Both can be displayed in geographical space. Three of the designs exemplified in Fig. 3 (road-based, stratified random by altitude, and model-based design) lead to very low but non-null probabilities, and to distinctly different probability patterns. This has large consequences for the final parameter estimates: similar data sets resulting from two different sampling designs will lead to different inferences because the weights given to individual observations are design-dependent (Lemeshow and Levy 1999).

#### *On the estimates of simple and complex parameters*

Our simulation study illustrates that the efficiency of the sampling design depends on the question and the parameters to be estimated (e.g. mean vs optimum, Fig. 4). If the basic aim is to estimate the mean value of a biological variable over a study area, the simple random sampling design, one of the simplest probabilistic designs, gives the results closest to truth. In contrast, if the basic aim is to estimate the optimum of the biological variable, then the model-based design gives the results closest to truth, even if the model used to estimate the optimum was simpler than

the one used to construct the simulation study. It is not surprising that the model-based design performed better than the other methods to estimate models, as it represents an iterative learning method. The so-called surface response methodology, which is derived from model-based sampling, was developed to estimate the kind of response shapes used in our virtual example (Box and Hunter 1957, Wu and Hamada 2000).

#### ***On the required sampling effort***

A key result from this semi-virtual study is that sample size was not the most influential factor in our ability to draw biological inference; indeed, sampling design had a stronger influence on the precision of parameter estimation than did sample size. This means that, concerning parameter estimation, appropriate design could reduce the sampling effort and costs associated with sampling, an important attribute given that resources for many ecological investigations are generally limited. It suggests, for example, that sub-sampling large databases with a design for deriving a dataset for model building may be preferred over using the whole dataset if the original design is unknown.

#### **Using prior knowledge of the study system**

Our simulation study also illustrates the importance of a priori knowledge of the system when establishing a sampling design. Basing sampling designs on prior knowledge and detailed hypotheses of how the main variables of the study system will impact the variables of interest is useful for reducing bias and increasing precision in the estimations of parameters with regard to the research question and to the costs. Incorporating prior knowledge of the study system is exactly the principle applied in Bayesian statistics in order to make reliable statistical inferences. Prior knowledge is, however, too often ignored and typically unknown, which makes it a challenging task (Clark 2005). There are potentially three ways to generate an efficient sampling design prior to starting field work. First, one can use expert knowledge and past studies to identify the research question, variables to measure, the corresponding target population, and from this derive an appropriate sampling frame. Expert knowledge is necessary but insufficient as a sole attribute of design, as it usually ignores other important variables or gradients that are not yet identified as important.

Second, one can perform a preliminary descriptive analysis in order to develop a design tailored to the question and study system. A very simple sampling scheme (e.g. simple random sample) can be first conducted on the variable of interest and then the response between the environmental and biological variable can be evaluated. A derived, more complex sampling design can then be based on the first results, such as surface response methodology (Box and Hunter 1957, Wu and Hamada 2000), from which we derived our model-based design. This powerful methodology effectively captures variation along a gradient and relies on a priori knowledge or assumptions on the patterns of interest, but is rarely used in ecology (but see Inouye 2001). Response surface methodology is particularly efficient to investigate effects of multiple factors

simultaneously (Inouye 2005). Using previously known distributions (Brito et al. 1999, Albert et al. 2010) or habitat suitability models (Guisan et al. 2006, Singh et al. 2009) as prior knowledge to establish or improve the sampling design has rarely been applied in ecology (but see Edwards et al. 2005).

Third, one can develop simulation studies, that are important tools for investigating ecological processes (Hirzel and Guisan 2002, Austin et al. 2006, Zurell et al. 2010), to aid in the development of appropriate sampling designs. We feel that such pathways have not been sufficiently exploited in ecology studies (but see Dengler and Oldeland 2010). Using real data and expert knowledge, as used in our simulation study, one can: 1) construct acceptable hypotheses on the expected patterns; 2) simulate different sampling designs including, for instance, the costs of sampling or the difficulty to reach the sampling sites; and 3) test the effect of sampling design (distribution of the sample in the sampling space) and effort on estimates (e.g. number of populations, number of individuals per populations). Following these steps can support decision-making, aid in developing sampling procedures better adapted to the research question, the hypotheses and the constraints of the study, and in particular it may enhance research flexibility (spare solutions) when facing the complexities that are inherent to field work.

#### **Estimating response curves, models or derived parameters**

Many ecological questions rely on accurate estimations of response curves and surfaces (Austin 1987). Moreover, the selection of models by balancing information content with parsimony is a crucial question in model estimation (Hastie et al. 2009). Many ecologically relevant parameters can be derived from models, such as optima (e.g. trait or growth optimal along an environmental gradient, Albert et al. 2010), breadth (e.g. niche breadth along an environmental gradient, Thuiller et al. 2004), or slope (e.g. rate of change in time or space). However, these aspects have not been well-linked to the sampling design, although we showed the importance of these issues when estimating different types of parameters (simple vs complex).

In particular, the generalization of the Horvitz-Thompson weighting of observations is not straightforward for the estimation of nonlinear parameters (e.g. the optimum of a quadratic curve, as in the simulation study). Estimation of nonlinear statistics from complex surveys has long been considered a difficult problem in the sampling literature, and different approaches have been suggested (Kish and Frankel 1974, Campbell 1980, Deville 1999, Berger and Skinner 2005), based either on a linearization of the estimator (so that the classical Horvitz-Thompson approach can be used) or on resampling (jackknife or bootstrap, Shao 1996). To the best of our knowledge, these approaches have not been routinely applied in ecology and evolution.

Moreover, the way model should be selected (Hastie et al. 2009) by balancing information with parsimony is a crucial point for the question of model estimation. This is not addressed sufficiently in the sampling literature but goes beyond the goal of our study. In sum, we believe that there



is a need to assess how to estimate ecologically relevant parameters which are both unbiased and nonlinear, and how to evaluate such estimates both in terms of specific designs (e.g. surface response methodology) and post-sampling analysis (e.g. Horvitz-Thompson approach) to address important questions in ecology and evolution.

*Acknowledgements* – The idea for this manuscript arose at a workshop “The Utility of Species Distribution Models as Tools for Assessing Impacts of Global Change” at Riederalp, Switzerland (August 2008). CA and WT acknowledge support from the ANR DIVERSITALP (ANR-07-BDIV-014) and EVORANGE (ANR-08-PEXT-11) projects. WT, NGY, TCE and NEZ received support from European Commission’s FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, No. 066866 GOCE) project. We thank Gretchen Moisen, David W. Roberts, Peter Pearman and one anonymous referee for very helpful and constructive comments on earlier drafts of the manuscript. Mention of a product by name does not constitute endorsement by the U.S. Geological Society.

## References

- Albert, C. H. et al. 2010. Intraspecific functional variability: extent, structure and sources of variation. – *J. Ecol.* 98: 604–613.
- Araujo, M. B. et al. 2009. Reopening the climate envelope reveals macroscale associations with climate in European birds. – *Proc. Natl Acad. Sci. USA* 106: E45–E46.
- Austin, M. P. 1987. Models for the analysis of species’ response to environment gradients. – *Vegetatio* 69: 35–45.
- Austin, M. P. and Smith, T. M. 1989. A new model for the continuum concept. – *Vegetatio* 83: 35–47.
- Austin, M. P. et al. 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. – *Ecol. Model.* 199: 197–216.
- Bart, J. et al. 1995. Reliability of the breeding bird survey: effects of restricting surveys to roads. – *Auk* 112: 758–761.
- Benichou, P. and Le Breton, O. 1987. Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. – *La Météorologie* 7: 23–24.
- Berger, Y. G. and Skinner, C. J. 2005. A jackknife variance estimator for unequal probability sampling. – *J. R. Stat. Soc. B* 67: 79–89.
- Box, G. E. P. and Hunter, J. S. 1957. Multi-factor experimental designs for exploring response surfaces. – *Ann. Math. Stat.* 28: 195–241.
- Brito, J. C. et al. 1999. Modelling wildlife distributions: logistic multiple regression vs overlap analysis. – *Ecography* 22: 251–260.
- Campbell, C. 1980. A different view of finite population estimation. – In: *Proc. Surv. Res. Meth. Sect. Am. Stat. Assoc.*, pp. 319–324.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. – *Ecol. Lett.* 8: 2–14.
- Courbois, J.-Y. et al. 2008. Evaluating probability sampling strategies for estimating redd counts: an example with chinook salmon (*Oncorhynchus tshawytscha*). – *Can. J. Fish. Aquat. Sci.* 65: 1814–1830.
- Dengler, J. and Oldeland, J. 2010. Effects of sampling protocol on the shapes of species richness curves. – *J. Biogeogr.*, in press.
- Deville, J.-C. 1999. Variance estimation for complex statistics and estimators: linearization and residual techniques. – *Surv. Methodol.* 25: 193–203.
- Edwards, T. C. et al. 2004. Assessing rarity of species with low detectability: lichens in Pacific Northwest forests. – *Ecol. Appl.* 14: 414–424.
- Edwards, T. C. et al. 2005. Model-based stratifications for enhancing the detection of rare ecological events. – *Ecology* 86: 1081–1090.
- Edwards, T. C. et al. 2006. Effects of underlying sample survey designs on the utility of classification tree models in ecology. – *Ecol. Model.* 199: 132–141.
- Gaston, K. J. 2003. The structure and dynamics of geographic ranges. – Oxford Univ. Press.
- Guisan, A. et al. 2006. Using niche-based models to improve the sampling of rare species. – *Conserv. Biol.* 20: 501–511.
- Haddad, N. M. et al. 2008. Determining optimal population monitoring for rare butterflies. – *Conserv. Biol.* 22: 929–940.
- Hastie, T. et al. 2009. The elements of statistical learning. Data mining, inference and prediction. – Springer.
- Hirzel, A. H. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling. – *Ecol. Model.* 157: 331–341.
- Hughes, R. M. and Peck, D. V. 2008. Acquiring data for large aquatic resource surveys: the art of compromise among species, logistics, and reality. – *J. North Am. Benthol. Soc.* 27: 837–859.
- Hutchinson, G. E. 1957. Concluding remarks. – Cold Spring Harbor Symp. Quant. Biol. 22: 145–159.
- Inouye, B. D. 2001. Response surface experimental designs for investigating interspecific competition. – *Ecology* 82: 2696–2706.
- Inouye, B. D. 2005. The importance of the variance around the mean effect size of ecological processes: comment. – *Ecology* 86: 262–265.
- Kish, L. and Frankel, M. R. 1974. Inference from complex samples. – *J. R. Stat. Soc. B* 36: 1–37.
- Körner, C. 2007. The use of “altitude” in ecological research. – *Trends Ecol Evol.* 22: 570–574.
- Lemeshow, S. and Levy, P. S. 1999. Sampling of populations: methods and applications. – Wiley.
- Mackenzie, D. I. 2006. Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. – *J. Wildl. Manage.* 70: 368–374.
- Magnani, F. 2009. Phenotypic variability: underlying mechanisms and limits do matter. – *New Phytol.* 184: 277–279.
- McCain, C. M. 2009. Vertebrate range sizes indicate that mountains may be ‘higher’ in the tropics. – *Ecol. Lett.* 12: 550–560.
- Mohler, C. L. 1983. Effect of sampling pattern on estimation of species distributions along gradients. – *Vegetatio* 54: 97–102.
- Olsen, A. R. et al. 1999. Statistical issues for monitoring ecological and natural resources in the United States. – *Environ. Monit. Assess.* 54: 1–45.
- Peterjohn, B. G. 1994. The North American breeding bird survey. – *J. Am. Birding Assoc.* 26: 386–398.
- R Development Core Team 2008. R: a language and environment for statistical computing. – R Foundation for Statistical Computing.
- Rosenstock, S. S. et al. 2002. Landbird counting techniques: current practices and an alternative. – *Auk* 119: 46–53.
- Särndal, G.-E. et al. 1992. Model assisted survey sampling. – Springer.
- Sauer, J. R. et al. 2003. Use of North American breeding bird survey data to estimate population change for bird conservation regions. – *J. Wildl. Manage.* 37: 372–389.
- Schreuder, H. T. et al. 1999. Relative costs and benefits of a continuous and a periodic forest inventory in Minnesota. – *Environ. Monit. Assess.* 59: 135–144.

- Scott, E. M. et al. 2008. Choice and criteria for selection of sampling strategies in environmental radioactivity monitoring. – *Appl. Radiat. Isot.* 66: 1575–1581.
- Shao, J. 1996. Resampling methods in sample surveys. – *Statistics* 27: 203–254.
- Singh, N. J. et al. 2009. Using habitat suitability models to sample rare species in high-altitude ecosystems: a case study with Tibetan argali. – *Biodivers. Conserv.* 18: 2893–2908.
- Skalski, J. R. et al. 2005. Wildlife demography: analysis of sex, age and count data. – Elsevier.
- Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – *Ecography* 27: 165–172.
- Thuiller, W. et al. 2010. Variation in habitat suitability models does not always relate to variation in species' plant functional traits. – *Biol. Lett.* 6: 120–123.
- Van Horne, B. 2002. Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. – In: Scott, J. M. et al. (eds), *Predicting species occurrences: issues of accuracy and scale*. Island Press, pp. 63–72.
- Wintle, B. A. et al. 2005. Estimating and dealing with detectability in occupancy surveys for forest owls and arboreal marsupials. – *J. Wildl. Manage.* 69: 905–917.
- Wu, C. F. J. and Hamada, M. 2000. *Experiments: planning, analysis, and parameter design optimization*. – Wiley.
- Yoccoz, N. G. et al. 2001. Monitoring of biological diversity in space and time. – *Trends Ecol. Evol.* 16: 446–453.
- Zurell, D. et al. 2010. The virtual ecologist approach: simulating data and observers. – *Oikos* 119: 622–635.