Authors: Silvio Marta[1,2,*], Federica Lacasella[3], Antonio Romano[4,5], Gentile Francesco Ficetola[1,6]

Title: Cost-effective spatial sampling designs for field surveys of species distribution

Research addresses:

[1] Department of Environmental Science and Policy, University of Milan, Via Celoria 26, 20133 Milan, Italy

[2] Institute of Ecosystem Studies - National Research Council c/o Department of Biology, University of Rome “Tor Vergata”, Via della Ricerca Scientifica, 00133 Rome, Italy

[3] Department of Entomology, University of Wisconsin, 1552 University Avenue - Wisconsin Energy Institute, 53706 Madison, WI, USA

[4] Istituto per i Sistemi Agricoli e Forestali del Mediterraneo - National Research Council of Italy, Via Patacca 85, 80056, Ercolano (NA), Italy

[5] MUSE - Museo delle Scienze, Corso del Lavoro e della Scienza 3, 38122 Trento, Italy

[6] Université Grenoble Alpes, CNRS, Laboratoire d’Écologie Alpine (LECA), 38000 Grenoble, France

Corresponding Author (*): Silvio Marta - Via Celoria 26, 20133 Milan, Italy - e-mail: silvio.marta@hotmail.it - mobile: +39 3779625871

ORCID:

Silvio Marta: https://orcid.org/0000-0001-8850-610X

Antonio Romano: https://orcid.org/0000-0002-5029-4372

Gentile Francesco Ficetola: https://orcid.org/0000-0003-3414-5155

Abstract

High-quality biodiversity inventories are key tools to develop effective conservation strategies, but financial resources devoted to systematic species inventories are usually limited. Different sampling strategies have been proposed to efficiently allocate such limited resources (i.e. accessibility-based, stratified random and grid samplings), but their effectiveness may depend on the aim of the survey. Our aim was to assess which approach can provide the best trade-off between sampling costs and accuracy in estimating both single species distribution and regional species set composition. We generated simulated species distribution data to compare costs and performances of the three sampling methods in assessing species distribution. When we aim at measuring species range (i.e. area of occupancy or extent of occurrence), or obtaining baseline ecological data for conservation assessments (i.e. niche breadth), grid sampling usually provided the best trade-off between performances and costs at both the species and regional levels. Otherwise, the stratified random sampling outperformed the other methods when we are interested in assessing the relative rarity (i.e. species frequency) of the species across the study area. Low quality distribution data can lead to heavily biased conclusions on biodiversity trends or impacts of environmental changes; our findings highlight that selecting the right sampling strategy is essential to obtain reliable estimates of both single species distribution and regional species set composition.

Keywords: field survey design, grid sampling, stratified random sampling, accessibility-based sampling, species frequencies, area of occupancy, extent of occurrence, niche breadth

Introduction

Species inventories are a key tool to obtain baseline data on the distribution of organisms and to develop effective conservation strategies (Barthlott & Winiger 1998). Systematic field surveys can enhance our knowledge of species occurrences and relative frequencies, which are essential to detect and track changes in biodiversity patterns (e.g. modifications in species richness or community composition following climate change, urbanization or agricultural intensification), to identify species or areas of high conservation priority, and to develop successful management measures (Austin & Heyligers 1989; Neldner et al. 1995; Hortal & Lobo 2005). Although survey campaigns are widely acknowledged as a primary tool in conservation planning and management, human and financial resources devoted to biodiversity survey and monitoring are limited. As a consequence, one of the main issues for conservationists and managers remains how to allocate limited resources to carry out the best conservation outcomes (McCarthy et al. 2012; Ficetola et al. 2018).

Surveying costs, in terms of time and/or funds, can be reduced by selecting sampling sites that are more easily accessible, usually close to roads ("accessibility-based" sampling) (Greenwood, 1996; Jobe & White 2009). However, site accessibility is seldom uniform across a region. For instance, road distribution is related to multiple factors, such as the physical properties of the landscape (e.g. elevation, orography, presence of barriers), and the distribution of human activities (e.g. presence of urban, agricultural or industrial areas) (Nelson 2008; Uchida & Nelson 2010). Therefore, easily accessible sites are often associated with anthropogenic stresses that are likely to affect species distribution. Many plant and animal species show limited frequency and / or activity nearby roads (e.g. edge effect) because of lower habitat quality and increased mortality (Forman & Alexander 1998; Trombulak & Frissell 2000; Fahrig & Rytwinski 2009). As a consequence, even if appealing from a cost perspective, accessibility-based samplings may provide spatially and/or ecologically biased data

(Kadmon et al. 2004). It is thus fundamental that these aspects are carefully accounted for before any inference is made about patterns and potential drivers of biodiversity.

Given the spatial bias of many species distribution datasets (Ficetola et al. 2013; Yang et al. 2014), several methods have been proposed to optimize and standardize efforts in collecting biodiversity information across a given area. Stratified (habitat-specific) random and grid sampling are among the most popular methods (Smith et al. 2017). However, outputs, spatial bias and costs may be very different among these methods, and their effectiveness mostly depends on the aims of the study. Stratified random sampling could return spatially unbiased information about species distribution and frequency across the study area by sampling all the potential suitable habitats (Yoccoz et al. 2001; Smith et al. 2017) but, due to logistic constraints, its application may be limited to surveying a reduced number of taxa in relatively small study areas (Guisan & Zimmermann 2000). This method seems particularly appropriate for investigating the distribution of rare or endangered species with well-known ecological constraints, as it requires some *a-priori* knowledge of the requirements of target species (e.g. inhabited vegetation types, elevational range); consequently, setting up a multi-habitat and multi-species (i.e. assemblage level) stratified sampling over large study areas can be technically complex and expensive (Guisan & Zimmermann 2000). Grid sampling (systematic survey *sensu* Wessels et al. 1998) could be more appropriate if the aim is to collect data on distribution patterns on a large set of species (e.g. assemblages) within a study area. In this case, a uniform sampling of the study area would be desirable. This approach could provide spatially unbiased estimates of species distribution, which are helpful to map biodiversity patterns within the study area; however it could be excessively expensive, and may not always lead to reliable estimates of species frequencies (Overton & Lehmann 2003). Even if statistically representative, both of these approaches may nevertheless under-represent or even lack species living in extremely rare habitats, for which ad-hoc strategies of site selection could be advisable (Økland 2007; Rolaček et al. 2007).

The choice of the sampling method is a crucial and challenging task that requires awareness about the strengths and weaknesses associated with each sampling approach. The relative performances and costs of different approaches may be assessed by comparing data collected with different protocols in the same area (Kadmon et al. 2004; Mccarthy et al. 2012). However, no method provides a perfect knowledge of true species distribution, thus hampering the estimation of the absolute biases. The analysis of simulated data on species distribution provides several advantages, such as the perfect knowledge of species occupancy and frequency, and community composition across the study area; this, in turn, allows the quantification of the sampling bias in relation to the real pattern (i.e. the "truth"), and the comparison of the biases of estimators based on different sampling methods (Hirzel & Guisan 2002; Zurell et al. 2010; Smith et al. 2017).

Here we used simulated species distribution data to compare costs (in terms of time needed to reach and survey the sites; i.e. total time) and performances of three different sampling methods (accessibility-based, stratified random and grid samplings) in assessing both single species distribution and species set composition across the study area. Stratified random and grid are rigorous sampling strategies, which can allow unbiased estimation of the parameters of interest (Smith et al. 2017). On the contrary, accessibility-based sampling often has high bias, but such data are frequent in occasional inventories, thus it is important to assess their relative performance. We considered three landscapes configurations, differing for their accessibility (i.e. road densities) and also assessed the robustness of our results to the issues of imperfect detection (MacKenzie et al. 2006; Kery & Royle 2016) and edge effect (Palomino & Carrascal 2007; Semlitsch et al. 2007), given their pervasive effects on species distribution data and on the reliability of survey results. Water dependent organisms were selected as it is easy to identify relationships between the distribution of presence sites (i.e. waterbodies) and accessibility, but results can be applied to many organisms that can be sampled in sites where appropriate resources (habitats) are. The aim of our study was to provide guidelines for researchers as well as for non-profit organization and government agencies dealing with biodiversity survey and monitoring. This will allow optimizing sampling design depending on both the survey aim and available resources, thus maximizing the reliability of the gathered data in term of species distribution.

Methods

Simulated species and landscape

Our simulation approach mimicked surveys aiming at detecting water-dependent organisms (e.g. amphibians, water birds, insects, or any kind of aquatic taxon). Artificial distribution data were generated for 15 hypothetical aquatic species differing in their habitat preferences, response to elevational gradients, and occupancy probabilities. For habitat preferences, we considered three species typologies: specialists for lentic habitats (e.g. ponds or small lakes), specialist of lotic habitats (e.g. streams), generalist (present in both typologies; Table 1). For elevation, each species showed an optimal elevation, and we assumed a Gaussian response to the altitudinal gradients (i.e. each species responded to the elevational variation with a symmetrical and decreasing occurrence probability around an optimum value, following a Gaussian probability curve). Species differed in optimum value (mean) and amplitude of their responses (standard deviation, sd) (see Table 1). Although variables other than elevation (e.g., water depth) also affect the distribution of aquatic species, and elevation may not be the key environmental driver of distribution *per se*, elevation is directly or indirectly linked to major variables (e.g. temperature, solar radiation, oxygen pressure, hydroperiod and wind), that can deeply influence organisms occurrence and frequency and overall biodiversity patterns (Guisan & Zimmermann 2000; Körner 2007; Graham et al. 2014). Furthermore, orography strongly determines the distribution of roads. To obtain realistic species distributions, occupancy probability was set to 0.5 (6 species) or 0.25 (9 species): only a randomly selected portion of suitable sites was thus considered effectively populated. Consequently, for each species, realized occupancy was higher around the optimum value (mean) and decreased following a Gaussian probability curve. Potential biotic interactions among simulated species were not considered. See Electronic Supplementary Material 1 (ESM1) for an example of the scripts used to generate species distribution data.

To obtain simulations mimicking the complexity of real landscapes, simulated data were generated on a true area of 40 × 40 km placed at the foothills of the Eastern Italian Alps (upperleft corner: x = 714,000 m, y = 5,114,000; Map projection: UTM zone 32N), characterized by an elevational range of more than 2,000 m. Patterns of spatial aggregation of lentic waters and paths of both roads and lotic waters are mainly determined by local orography, geomorphological and lithological features. Selecting a true area allowed us obtaining a realistic distribution of both sampling sites and road network without compromising the generality of results (Hirzel et al. 2001; Meynard & Quinn 2007).

Environmental variables

For the study area, elevation data were obtained from the Shuttle radar topographic mission (SRTM; original resolution = 3 arc-seconds; downloaded on 20th April 2010), reprojected to UTM 32N (resolution = 92.66 m) and slightly rescaled to vary between 0 and 2,252 m a.s.l. (Figure 1a). The complete road network was obtained from the database DBPrior10K (downloaded on 15th January 2016 from http://www.centrointerregionale-gis.it/DBPrior/DBPrior.asp). Single roads, both main and secondary roads (branches), were manually reclassified to three different classes (class 1 to class 3; Figure 1b). In our simulations we explored three scenarios of accessibility (low, medium and high road densities). In the low accessibility scenario we only considered class 1 roads (main roads); class 1 + 2 roads (main roads and their first branches) were considered in the medium accessibility scenario, and for the high accessibility scenario we considered all roads as exploitable during the survey.

Sampling sites included both lentic and lotic sites. Lotic sites were obtained by simplifying the hydrographic network available on the Italian National Geoportal website (downloaded on 7th November 2015 from http://wms.pcn.minambiente.it/ogc?map=/ms_ogc/wfs/Aste_fluviali.map via the Web Feature Service in Quantum GIS 2.2). For each stream we set a sampling site every 1,500 m with a minimum of 2 sampling sites per stream, obtaining a total of 719 lotic sites (Figure 1c). Lentic sites were detected from the toponym layer (downloaded on 13th November 2015 from http://wms.pcn.minambiente.it/ogc?map=/ms_ogc/wfs/Toponimi_2011.map via the Web Feature Service in Quantum GIS 2.2), by selecting sites representing water-related typologies (118 points). Available maps certainly underestimate lentic sites, given that small ponds are often undetected by aerial photos (Ficetola et al. 2015). To approximate a 2:1 ratio between lotic and lentic sampling sites and retain at the same time the spatial aggregation pattern typical of lentic habitats, we randomly generated 225 additional lentic points within a buffer of 2,000 m from the extant ones (total lentic sites = 343; Figure 1c). This led to a total of 1,062 sampling sites (719 lotic + 343 lentic sites). For each potential sampling site, travelling costs (in term of time) were calculated using the gdistance R package (van Etten 2015) and applying the Tobler's Hiking Function. This function provides a rough estimate for the maximum speed of off-path hiking given the slope of the terrain (Tobler 1993). Once obtained the inter-cell speed (m/s), the correction (ratio) for the inter-centroid distance converts the speeds in reciprocal of times (1/s): simply summing the reciprocal of these reciprocals (Σ 1/(1/s)) allow us to obtain the total travelling time. For each of the three accessibility scenarios, costs were estimated between each sampling site and the closest road. Despite in the real world it is not always feasible to gain access to the whole set of sampling sites, here we considered all sites potentially accessible and differing only in the travelling cost to be spent in reaching them.

Survey design

We evaluated three survey strategies (grid, stratified random and accessibility-based samplings) under three scenarios of accessibility (low, medium and high). In 999 simulations, we generated the distribution of artificial species; simulated species sets were then sampled according to the three different methods (see Supplementary Figure 1b-d in ESM2 for an example of site selection). To simplify comparisons, we employed the same sampling effort (i.e. same number of sampling sites) in the three sampling methods. Grid sampling was performed by building grids of different cell size and selecting, whenever present, one lotic and one lentic site within each cell of the grid. To account for scale dependent effects, analyses were run using cell sizes of 10, 6.67, 5, 4, 3.33 and 2.5 km (corresponding to 32, 69, 118, 167, 235 and 373 sampling sites). We applied the same sampling effort to the three methods, thus the same number of sampling sites (*n*) used in the grid approach was subsequently sampled with the stratified random and accessibility-based methods. For the stratified random sampling we considered just one ecologically informative stratum, i.e. the availability of water resources (both streams and ponds) across the whole study area. Sampling was then performed by randomly selecting from the whole dataset of water resources *n* sampling sites. Only for the accessibility-based sampling, we selected the *n* sampling sites with the lowest travelling costs; consequently, the total cost is the same for all the replicates with the same *n* within the same accessibility scenario. Travelling cost estimation and sampling selection were repeated for each of three accessibility scenarios. For purpose of comparison, two additional values of *n* (600 and 750 sites) were further sampled with the accessibility-based sampling only. A total of 60 combinations were thus analysed for each of the 999 simulated species sets: 3 sampling methods × 6 sampling efforts × 3 accessibility scenarios, plus two additional sampling efforts (i.e. 600 and 750 sites) × 3 accessibility scenarios for the accessibility-based sampling only.

We performed two additional simulation runs to assess the impact of edge effect and imperfect detection on our conclusions. To assess the consequences of edge effect, sites within 90 m from roads were considered unsuitable for the target species (average travel time: about 110 s from the nearest road), all other parameters being constant. Furthermore, in standard analyses, we assumed just one survey per site and perfect detection of all the present species. However, detection probability is almost always below one, and multiple surveys are needed to obtain robust estimates of species distribution (MacKenzie et al. 2006; Petitot et al. 2014). We therefore repeated simulations assuming that species have imperfect detection; the detection probability of each species was randomly drawn from the interval [0.1,0.7]. Each site was surveyed in three distinct sampling occasions, while all the other parameters remained consistent with the other simulations.

Assessing the efficiency of survey methods

The performance of each survey method (grid, stratified random and accessibility-based methods) was evaluated by its ability to assess species distribution at a given survey cost. At the regional level, two measures of species distribution were used, reflecting different survey aims: area of occupancy and species frequency across the landscape. Area of occupancy is a measure of the spatial distribution of species, while frequency across the landscape is the proportion of sites with species presence. These two metrics are not necessarily correlated and allow to describe and represent different forms of rarity (Rabinowitz 1981). For instance, a species can occupy a very large number of sites within a small area (e.g. small range species that are locally abundant), or can occupy very large ranges with just a few populations (sparse populations over broad ranges). For each cell size used during the grid sampling (i.e. 10, 6.67, 5, 4, 3.33 and 2.5 km), area of occupancy was calculated as the total number of cells in which a given species was present (true occupancy) or collected (sampled occupancy) standardized by total number of cells; this approach is similar to the one used during IUCN species assessment. Species frequency across the study area was calculated as the total number of sites in which the species was present (true frequency) or collected (sampled frequency), standardized by the total number of sites or the number of surveyed sites, respectively. At the regional level, bias was calculated as the overall Renkonen (Percentage) dissimilarity (Renkonen 1938) between standardized sampled (i.e. sampled occupancy or frequency) and true (i.e. true occupancy or frequency) species sets. Renkonen dissimilarity corresponds to Bray-Curtis dissimilarity when this is calculated on relative rather than absolute abundances, and solves the problem of density invariance highlighted for this latter index (Jost et al. 2011). At the species level, we measured performance also for two additional parameters: niche breadth and extent of occurrence. For each species, niche breadth was calculated as the altitudinal range experienced by the species, while extent of occurrence as the area contained within the minimum convex polygon enclosing all sites occupied by the species.

For each estimator of sampling performances, we calculated relative bias as (true value - estimated value) / true value. Consequently, bias values can range between -1 and 1 when dealing with species frequency across the landscape, and between 0 and 1 in all the other cases (i.e., area of occupancy, extent of occurrence and niche breadth). We report species-level measures of bias for a subset of species representing the whole range of simulated species: the commonest (Species 1), the rarest (Species 9) and one species with an intermediate frequency (Species 10).

In biodiversity surveys, the time required by operators to complete sampling is a major determinant of total survey cost. We used two metrics to measure the sampling cost of each survey scheme: cumulative travel time, and number of surveyed sites. Cumulative travel time was the sum of the time needed to reach all the n sites, as the time to reach survey sites constitute a major part of the working time of operators. Furthermore, we considered the total number of surveyed sites, as sampling more sites requires a larger effort. The number of surveyed sites ranged between 32 and 373 (up to 750 for the accessibility-based sampling only). These measures were calculated for each of the three different accessibility scenarios (from low to high road density). We finally calculated the total survey time as (number of surveyed sites × site sampling time) + cumulative travel times, by assuming an average sampling time of 20 minutes per site, which is a typical survey effort for the national monitoring of amphibians and reptiles in Italy (Stoch & Genovesi 2016). Times other than off-path hiking (e.g., driving time from a "base") and costs for materials (e.g., sampling equipment or fuel) were not considered for the calculation of costs as they strongly depend on the positioning of the base and the sampling methodology, respectively.

Analyses were performed using the R programming environment (R 3.2.5; R Development Core Team 2016) and associated packages (Goslee & Urban 2007; Bivand & Rundel 2014; Bivand et al. 2015; Hijmans 2015; van Etten 2015). Data sets and R scripts used to run the analyses are available as supplementary material (ESM1).

Results

Analyses of relationships between sampling costs and bias showed that an increase in total survey time was always associated with a decrease in sampling bias (Figs. 2-4). However the different monitoring strategies showed substantial differences in bias for all the measures of species distribution used, i.e. area of occupancy (Figs. 2a-c and 3a-c), frequency (Figs. 2d-f and 3d-f), extent of occurrence (Fig. 4a-c) or niche breadth (Fig. 4d-f), and across the accessibility scenario considered.

Regional level analysis: species area of occupancy

When we considered the reliability of estimates of area of occupancy across the whole species set and study area, the accessibility-based sampling always showed smaller total and travel times than the other methods (Fig. 2a-c and Supplementary Figure 2a-c in ESM2, respectively). The relative performances (biases) of the three methods considerably varied depending on the accessibility scenario (Fig. 2a-c). Grid sampling consistently provided the best estimates across all the accessibility scenarios, although accessibility-based samplings slightly outperformed the others when the greatest number of sites was sampled (750 sites). Accessibility-based and random sampling showed similar performance in the high and medium accessibility scenarios (Fig. 2a-b), while random sampling generally showed lower bias than the accessibility approach in the low accessibility one (Fig. 2c). See Supplementary Figure 2 in ESM2 for an estimation of sampling cost, separately showing cumulative travel times and number of sampled sites and Supplementary Figures 5a-c in ESM2 for the consequences that edge effect has on sampling bias.

Regional level analysis: species frequency in the landscape

When we considered the reliability of species frequency estimates across the whole species set and study area, the relative performances of each method were consistent across the three accessibility scenarios (Fig. 2d-f, Supplementary Figure 3 in ESM2). Stratified random sampling returned the most accurate estimation of the species set at the regional level (Fig. 2d-f), while the accessibility-based sampling provided the worst estimates, irrespective of the landscape accessibility and the measures of cost used. See Supplementary Figure 3 in ESM2 for an estimation of sampling cost, showing separately cumulative travel times and number of sampled sites and Supplementary Figures 5d-f in ESM2 for the consequences that edge effect has on sampling bias.

Species level analysis

The performances of the three sampling methods in describing area of occupancy, frequency, extent of occurrence and niche breadth of single species revealed patterns partially similar to the ones from the regional level analyses (Figs. 3 and 4). Here we focus on the results of the high accessibility scenario, but conclusions for the other scenarios were similar (Supplementary Fig. 4 in ESM2). For all the species, sampling bias ranged more widely with respect to the regional level analyses. Considering the bias in estimating the area of occupancy (Fig. 3a-c), the accessibility-based method showed the best performances for common species only (Fig. 3a), whereas grid sampling outperformed accessibility-based sampling for rare species (Fig. 3c). For the estimation of species frequencies across the landscape (Fig. 3d-f), the results are consistent with the patterns observed at the regional level: the stratified random sampling provides bias values very close to zero for all the species and thus clearly outperformed the other methodologies. For the estimation of the extent of occurrence (Fig. 4a-c), the accessibility-based sampling slightly outperformed the other methods for species with high and intermediate frequencies (Fig. 4a-b), while grid sampling showed the lowest bias when dealing with rare species (Fig. 4c). Lastly, considering the bias in estimating niche breadth (Fig. 4d-f) all the methods provide a similar performance for species with high and intermediate frequency (Fig. 4d-e), while grid sampling returned the best estimates for rare species (Fig. 4f).

Imperfect detection

At the regional level, the overall performances of the three sampling methods were consistent with previous results, when imperfect detection was included in simulations (Supplementary Figure 6 in ESM2). Grid and stratified random samplings returned the best estimates of area of occupancy (Fig. S6a-c) and species frequency (Fig. S6d-f), respectively, but incomplete detection and multiple sampling occasions increased both the uncertainties in estimating the species set at the regional level, and the sampling costs.

Discussion

Efficient and reliable biodiversity surveys are necessary to obtain distribution data, but substantial resources are required to obtain robust estimates of species range and frequency. At a given sampling cost, different approaches show strong heterogeneity in performance, and our results help to select the optimal sampling strategy depending on both the aims of the survey and the landscape accessibility.

When the main aim is obtaining measures of geographic range of species, baseline data for conservation assessments (IUCN 2001; Tracewski et al. 2016), or overall biodiversity patterns across the landscape, grid -based sampling provides a good trade-off between sampling bias and costs at both the regional and single species levels (Figs. 2a-c, 3a-c, 4). Accessibility-based sampling effectively estimated the area of occupancy of commonest species, but suffers multiple drawbacks. First, species distributions can be accessibility-biased (e.g. lower abundance nearby roads, a classical case of edge effect) (Palomino & Carrascal 2007; Semlitsch et al. 2007), and under these circumstances selecting sites on the basis of accessibility would provide biased results (discussed below). Furthermore, grid sampling considerably outperforms the accessibility-based one in estimating areas of occupancy level (Fig. 2) and the distribution of rare species (Fig. 3c, Fig. 4c and Fig. 4f). Grid sampling allows a homogeneous spatial distribution of sampling sites across the whole study area, thus providing more balanced estimates of single species relative distribution and maximising spatial coverage, which is essential for the assessment of species ranges. The grid approach we used can be particularly effective, as it may be seen as a grid-based stratified sampling: in fact, within each cell, two different typologies of sites (i.e. one lentic plus one lotic habitat) were randomly selected, allowing to take into account ecological variation and thus improving the overall quality of the estimates.

Conversely, if the main aim of the survey is to collect reliable data on species frequency across the landscape, the stratified random sampling outperformed the other methods in describing both regional patterns and single species frequencies (Fig. 2d-f, Fig. 3d-f). This can be due to its ability to gather data proportionally to the resource typology and spatial availability, allowing a more reliable estimation of species frequency within the study area. The excellent performance of random sampling in estimating species frequency at both the regional and the single species level was independent of landscape accessibility and the measures of cost used (Fig. 2d-f, Fig. 3d-f, Supplementary Figures 3 & 4d).

Occasional samplings are often biased by accessibility. As occasional sampling is a main source of biodiversity distribution data, accessibility-based sampling is perhaps the most frequent strategy for the collection of distribution data, even though this is only seldom explicitly stated. For instance, citizen science provides a huge amount of data over large temporal and spatial scales but it is prone to spatial biases from infrastructure and human population density (Geldmann et al. 2016) because roads, cities, and other physical features determine accessibility for observers. This bias may be reduced using effective protocol development and volunteer training (Flesch & Belt 2017), still it remains pervasive in biodiversity datasets. In principle, selecting sampling sites on the basis of accessibility greatly reduces sampling time, and thus allows visiting a larger number of sites. For instance, in this study the travel time needed to visit the 373 most accessible sites (53 h) was about seven times lower than the time required to visit the same number of sites selected using the alternatives schemes (355 and 362 h for grid and stratified random sampling, respectively), in the intermediate accessibility landscape (Supplementary Fig. 2b). Unfortunately, surveying such a large number of sites does not improve the quality of results, confirming the existing concerns on road-biased sampling. Accessibility-based sampling is sometimes thought to represent the most cost-effective solution to sample an area (Albert et al. 2010), but its effectiveness strongly depends on the density of the road network: in fact, sampling sites close to roads reduces costs only within highly accessible landscapes or for common species (Fig. 2, Fig. 3a and Fig. 4a and d), and only if road distribution is not heavily biased by spatial and ecological features (e.g. landscape composition or orography). Given that such biases are widespread, and given that the usefulness of the accessibility-based sampling is restricted to specific conditions, if possible other sampling strategies should be preferred in most of programmes.

In addition, roads often have negative effects such as direct killing by vehicles, disturbance, barrier effects and pollution (Forman & Alexander 1998; Rytwinski & Fahrig 2015). Consequently, occupancy is generally reduced in sites nearby roads (edge effect) (Palomino & Carrascal 2007; Semlitsch et al. 2007), posing additional issues to the accessibility-based sampling. If we assume that sites within 90 m from roads are unsuitable for the target species, accessibility-based sampling becomes even less reliable (Supplementary Fig. 5 in ESM2). When we estimate area of occupancy and species frequencies accounting for edge effect, the performances of the accessibility-based survey were far from being reliable. In practice, edge effects determine the highest observed bias values (Supplementary Fig. 5), and completely erases any potential advantage of accessibility-based sampling. Nevertheless, the interactions between roads, species occurrence, accessibility, and performance of surveys can be complex, and there are cases in which performing sampling along roads do not provide biased estimates of species distribution (Mccarthy et al. 2012).

In the real world, imperfect detection of species is pervasive, further increasing the complexity of planning biodiversity surveys (MacKenzie et al. 2006; Petitot et al. 2014). If detection is imperfect, multiple visits must be performed to each site, thus increasing the overall cost and the uncertainties of species distribution estimates. Nevertheless, after taking into account imperfect detection we obtained the same overall pattern, with grid sampling providing the best assessment of species range, and stratified sampling providing the best assessments of species frequencies (Supplementary Figure 6 in ESM2). This is probably due to the fact that detection probability was not different among sites with different accessibility, and the number of surveys per site was adequate to obtain reliable estimates of species occupancy. The situation could be more challenging when detection probability of species is not spatially random (Gu & Swihart 2004). For instance, species detection might be lower for rare species (Tanadini & Schmidt 2011) or nearby roads: in this case we expect that non-random imperfect detection would further increase the bias of accessibility-based sampling.

Our simulations were developed assuming aquatic species as the target of the survey and testing the effectiveness of three alternative sampling strategies. Small wetlands and streams often are discrete habitats, thus an *a-priori* selection of sites with a stratified sampling can be easily performed using geographic information systems, if information on wetland distribution is available. The selection of sampling sites may be more complex for terrestrial or marine organisms, whose habitats are often represented as polygon-like features (Smith et al. 2017). For these organisms there is the additional question of the appropriate position of the sampling site, within the polygon extent. The definition of the appropriate sampling site (e.g. point count, transect or trapping station) is strongly dependent on the study taxon and on the research aims, and is beyond the scope of the present study. Still, the increasing availability of informative strata (e.g. habitat typology, altitude, and microclimate data layers) can allow integrating multiple information sources, in order to optimize the sampling strategy even in the most complex situations. Therefore, grid and stratified random sampling can also be used for the selection of sampling sites for terrestrial organisms, once the potential sampling sites have been defined. At the same time, alternative sampling strategies such as the generalized random-tessellation stratified (GRTS; Stevens & Olsen, 2004) and the gradient directed transects (grandsects; Gillison & Brewer, 1985; Wessels et al., 1998) could be just as reliable as those tested here to optimize and standardize efforts in collecting biodiversity information across a given area. All of these objective approaches to site selection have the advantage to strongly limit subjective choices driven by environmental attractiveness or accessibility (Soberón et al. 2000; Parnell et al. 2003; Moerman & Estabrook 2006; Romo et al. 2006).

There is not a single sampling approach suitable for all the circumstances and, when setting up a survey or monitoring programme, the optimal sampling strategy should be defined on the basis of the landscape structure and the aims of the programme (Yoccoz et al. 2001). If the aim is to collect unbiased data on the spatial distribution of the species (e.g. for a distribution atlas) and to use them to assess biodiversity patterns, a grid sampling, eventually associated with a stratified selection of sites within each cell, is the more appropriate and cost-effective method. Conversely, the stratified random sampling returns the best trade-off between data reliability and sampling cost, when the focus is on species frequencies (e.g. assessing species rarity). Monitoring programmes must be repeated in time, to discover potential biodiversity changes, assess the consequences of environmental modifications, and test whether populations are declining or increasing (Nichols & Williams 2006; Wintle et al. 2010; Ficetola et al. 2018). However, low quality distribution data can lead to heavily biased conclusions when we test species or biodiversity trends, and impacts of environmental changes (Yoccoz et al. 2001). Selecting an optimal and objective approach to survey or monitoring is important to optimize the results, but is also the key to obtain reliable assessments of the long-term trajectories of species and ecosystems, and thus to best inform conservation and management.

References

Albert CH, Yoccoz NG, Edwards Jr TC, Graham CH, Zimmermann NE, Thuiller W (2010) Sampling in ecology and evolution–bridging the gap between theory and practice. Ecography 33:1028-1037. https://doi.org/10.1111/j.1600-0587.2010.06421.x

Austin MP, Heyligers PC (1989) Vegetation survey design for conservation: gradsect sampling of forests in north-eastern New South Wales. Biol Conserv 50:13-32. https://doi.org/10.1016/0006-3207(89)90003-7

Barthlott W, Winiger M (1998) Biodiversity: a challenge for development research and policy. Springer Verlag, Berlin

Bivand R, Rundel C (2014) rgeos: Interface to Geometry Engine - Open Source (GEOS). R package version 0.3-8. URL https://CRAN.R-project.org/package=rgeos

Bivand R, Keitt T, Rowlingson B (2015) rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.9-3. URL https://CRAN.R-project.org/package=rgdal

van Etten J (2015) gdistance: Distances and Routes on Geographical Grids. R package version 1.1-9. URL https://CRAN.R-project.org/package=gdistance

Fahrig L, Rytwinski T (2009) Effects of roads on animal abundance: an empirical review and synthesis. Ecol Soc 14:21

Ficetola GF, Bonardi A, Sindaco R, Padoa-Schioppa E (2013) Estimating patterns of reptile biodiversity in remote regions. J Biogeogr 40:1202-1211. https://doi.org/10.1111/jbi.12060

Ficetola GF, Rondinini C, Bonardi A, Baisero D, Padoa-Schioppa E (2015) Habitat availability for amphibians and extinction threat: a global analysis. Divers Distrib 21:302-311. https://doi.org/10.1111/ddi.12296

Ficetola GF, Romano A, Salvidio S, Sindaco R (2018) Optimizing monitoring schemes to detect trends in abundance over broad scales. Anim Conserv 21:221-231. https://doi.org/10.1111/acv.12356

Flesch EP, Belt JJ (2017) Comparing citizen science and professional data to evaluate extrapolated mountain goat distribution models. Ecosphere. https://doi.org/10.1002/ecs2.1638

Forman RT, Alexander LE (1998) Roads and their major ecological effects. Ann Rev Ecol Syst 29:207-231. https://doi.org/10.1146/annurev.ecolsys.29.1.207

Geldmann J, Heilmann-Clausen J, Holm TE, Levinsky I, Markussen B, Olsen K, Rahbek C, Tøttrup AP (2016) What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. Divers Distrib 22:1139-1149. https://doi.org/10.1111/ddi.12477

Gillison AN, Brewer KRW (1985) The use of gradient directed transects or gradsects in natural resource surveys. J Environ Manage 20: 103-127.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22:1-19.

Graham CH, Carnaval AC, Cadena CD, Zamudio KR, Roberts TE, Parra JL, McCain CM, Bowie RCK, Craig Moritz C, Baines SB, Schneider CJ, VanDerWal J, Rahbek C, Kozak KH, Sanders NJ (2014) The origin and maintenance of montane diversity: integrating evolutionary and ecological processes. Ecography 37:711-719. https://doi.org/10.1111/ecog.00578

Greenwood JJD (1996) Basic techniques. In: Shuterland WJ (ed) Ecological census techniques: a handbook. Cambridge University Press, Cambridge, pp11-110.

Gu W, Swihart RK (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biol Conserv 116:195-203. https://doi.org/10.1016/S0006-3207(03)00190-3

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147-186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hijmans RJ (2015) raster: Geographic Data Analysis and Modeling. R package version 2.3-40. URL https://CRAN.R-project.org/package=raster.

Hirzel A, Helfer V, Metral F (2001) Assessing habitat-suitability models with a virtual species. Ecol Model 145:111-121. https://doi.org/10.1016/S0304-3800(01)00396-9

Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. Ecol Model 157:331-341. https://doi.org/10.1016/S0304-3800(02)00203-X

Hortal J, Lobo JM (2005) An ED-based protocol for optimal sampling of biodiversity. Biodivers Conserv 14:2913-2947. https://doi.org/10.1007/s10531-004-0224-z

IUCN (2001) IUCN Red List Categories and Criteria: Version 3.1. International Union for Conservation of Nature, Gland (CH) and Cambridge (UK).

Jobe RT, White PS (2009) A new cost-distance model for human accessibility and an evaluation of accessibility bias in permanent vegetation plots in Great Smoky Mountains National Park, USA. J Veg Sci 20:1099-1109. https://doi.org/10.1111/j.1654-1103.2009.01108.x

Jost L, Chao A, Chazdon RL (2011) Compositional similarity and β (beta) diversity. In: Magurran AE, McGill BJ (eds) Biological diversity: frontiers in measurement and assessment, Oxford University Press, Oxford, pp 66-84

Kadmon R, Farber O, Danin A (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. Ecol Appl 14:401-413. https://doi.org/10.1890/02-5364

Kery M, Royle JA (2015) Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press, Cambridge

Körner C (2007) The use of ‘altitude’ in ecological research. Trends Ecol Evol 22:569-574. https://doi.org/10.1016/j.tree.2007.09.006

MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey L, Hines JE (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, Cambridge

McCarthy DP, Donald PF, Scharlemann JPW, Buchanan GM, Balmford A, Green JMH, Bennun LA, Burgess ND, Fishpool LDC, Garnett ST, Leonard DL, Maloney RF, Morling P, Schaefer HM, Symes A, Wiedenfeld DA, Butchart SHM (2012) Financial costs of meeting global biodiversity conservation targets: current spending and unmet needs. Science 338:946-949. https://doi.org/10.1126/science.1229803

Mccarthy KP, Fletcher RJ, Rota CT, Hutto RL (2012) Predicting species distributions from samples collected along roadsides. Conserv Biol 26:68-77. https://doi.org/10.1111/j.1523-1739.2011.01754.x

Meynard CN, Quinn JF (2007) Predicting species distributions: a critical comparison of the most common statistical models using artificial species. J Biogeogr 34:1455-1469. https://doi.org/10.1111/j.1365-2699.2007.01720.x

Moerman DE, Estabrook GF (2006) The botanist effect: counties with maximal species richness tend to be home to universities and botanists. J Biogeogr 33:1969-1974. https://doi.org/10.1111/j.1365-2699.2006.01549.x

Neldner VJ, Crossley DC, Cofinas M (1995) Using geographic information systems (GIS) to determine the adequacy of sampling in vegetation surveys. Biol Conserv 73:1-17. https://doi.org/10.1016/0006-3207(95)90049-7

Nelson A (2008) Travel time to major cities: A global map of Accessibility. Global Environment Monitoring Unit - Joint Research Centre of the European Commission

Nichols JD, Williams BK (2006) Monitoring for conservation. Trends Ecol Evol 21:668-673. https://doi.org/10.1016/j.tree.2006.08.007

Økland RH (2007) Wise use of statistical tools in ecological field studies. Folia Geobot 42: 123-140. https://doi.org/10.1007/BF02893879

Overton JMcC, Lehmann A (2003) Predicting vegetation condition and weed distributions for systematic conservation management. New Zealand Department of Conservation

Palomino D, Carrascal LM (2007) Threshold distances to nearby cities and roads influence the bird community of a mosaic landscape. Biol Conserv 140:100-109. https://doi.org/10.1016/j.biocon.2007.07.029

Parnell JAN, Simpson DA, Moat J, Kirkup DW, Chantaranothai P, Boyce PC, Bygrave P, Dransfield S, Jebb MHP, Macklin J, Meade C, Middleton DJ, Muasya AM, Prajaksood A, Pendry CA, Pooma R, Suddee S, Wilkin P (2003) Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand. J Biogeogr 30:193-209. https://doi.org/10.1046/j.1365-2699.2003.00828.x

Petitot M, Manceau N, Geniez P, Besnard A (2014) Optimizing occupancy surveys by maximizing detection probability: application to amphibian monitoring in the Mediterranean region. Ecol Evol 4:3538-3549. https://doi.org/10.1002/ece3.1207

R Development Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. URL http://www.R-project.org

Rabinowitz D (1981) Seven forms of rarity. In: Synge H (ed) The biological aspects of rare plants conservation. John Wiley & Sons, Hoboken ,pp 205-217

Renkonen O (1938) Statistisch-ökologische Untersuchungen über die terrestrische Käferwelt der finnischen Bruchmoore. Ann Zool Soc Zool-Bot Fennicae 6:1-231.

Roleček J, Chytrý M, Hájek M, Lvončík S, Tichý L (2007) Sampling design in large-scale vegetation studies: Do not sacrifice ecological thinking to statistical purism!. Folia Geobot 42: 199-208. https://doi.org/10.1007/BF02893886

Romo H, García-Barros E, Lobo JM (2006) Identifying recorder-induced geographic bias in an Iberian butterfly database. Ecography 29: 873-885. https://doi.org/10.1111/j.2006.0906-7590.04680.x

Semlitsch RD, Ryan TJ, Hamed K, Chatfield M, Drehman B, Pekarek N, Spath M, Watland A (2007) Salamander abundance along road edges and within abandoned logging roads in Appalachian forests. Conserv Biol 21:159-167. https://doi.org/10.1111/j.1523-1739.2006.00571.x

Smith AN, Anderson MJ, Pawley MD (2017) Could ecologists be more random? Straightforward alternatives to haphazard spatial sampling. Ecography 40:1251-1255. https://doi.org/10.1111/ecog.02821

Soberón JM, Llorente JB, Oñate L (2000) The use of specimen-label databases for conservation purposes: an example using Mexican Papilionid and Pierid butterflies. Biodivers Conserv 9:1441-1466. https://doi.org/10.1023/A:1008987010383

Stevens DL Jr, Olsen AR (2004) Spatially balanced sampling of natural resources. J Am Stat Assoc, 99: 262-278. https://doi.org/10.1198/016214504000000250

Stoch F, Genovesi P(2016) Manuali per il monitoraggio di specie e habitat di interesse comunitario (Direttiva 92/43/CEE) in Italia: specie animali. Serie Manuali e linee guida, 141/2016, ISPRA

Tanadini LG, Schmidt BR (2011) Population size influences amphibian detection probability: implications for biodiversity monitoring programs. PLoS One. https://doi.org/10.1371/journal.pone.0028244

Tobler W (1993) Three presentations on geographical analysis and modelling. Technical report 93-1, NCGIA

Tracewski Ł, Butchart SHM, Di Marco M, Ficetola GF, Rondinini C, Symes A, Wheatley H, Beresford AE, Buchanan GM (2016) Toward quantification of the impact of 21st-century deforestation on the extinction risk of terrestrial vertebrates. Conserv Biol 30:1070-1079. https://doi.org/10.1111/cobi.12715

Trombulak SC, Frissell CA (2000) Review of ecological effects of roads on terrestrial and aquatic communities. *Conserv Biol* **14:**18-30 (2000). https://doi.org/10.1046/j.1523-1739.2000.99084.x

Uchida H, Nelson A (2010) Agglomeration Index: Towards a new measure of urban concentration. In: Beall J, Guha-Khasnobis B, Kanbur R (eds) Urbanization and development: Multidisciplinary perspectives. Oxford University Press, Oxford, pp 41-60

Wessels KJ, Van Jaarsveld AS, Grimbeek JD, Van der Linde MJ (1998) An evaluation of the gradsect biological survey method. Biodivers Conserv 7:1093-1121. https://doi.org/10.1023/A:1008899802456

Wintle BA, Runge MC, Bekessy SA (2010) Allocating monitoring effort in the face of unknown unknowns. Ecol Lett 13:1325-1337. https://doi.org/10.1111/j.1461-0248.2010.01514.x

Yang W, Ma K, Kreft H (2014) Environmental and socio-economic factors shaping the geography of floristic collections in China. Global Ecol Biogeogr 23:1284-1292. https://doi.org/10.1111/geb.12225

Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. Trends Ecol Evol 16:446-453. https://doi.org/10.1016/S0169-5347(01)02205-4

Zurell D, Berger U, Cabral JS, Jeltsch F, Meynard CN, Münkemüller T, Nehrbass N, Pagel J, Reineking B, Schröder B, Grimm V (2010) The virtual ecologist approach: simulating data and observers. Oikos, 119:622-635. https://doi.org/10.1111/j.1600-0706.2009.18284.x

Acknowledgements

Author Contributions

SM, AR and FL conceived the idea; SM and GFF designed the methodology; SM, GFF and FL analysed the data; FL and SM led the writing of the manuscript. All authors critically contributed to the drafts and gave final approval for publication.

Figure captions

Figure 1 (b/w in print)

Study area: Digital elevation model (a). Road network with classified roads: class 1 only = low accessibility scenario; class 1 + 2 = medium accessibility scenario; all classes = high accessibility scenario (b). Sampling stations (N = 1062) showing separately the 719 lotic sites (along streams) and the 343 lentic sites. Green triangles = lentic sites; blue circles = lotic sites (c).

Figure 2 (b/w in print)

Regional level: relationships between total sampling costs and bias for the three sampling methods (grid, random and accessibility-based samplings) at three different accessibility scenarios (high, medium and low road densities). Grid sampling was performed using cell sizes of 10, 6.67, 5, 4, 3.33 and 2.5 km (corresponding to 32, 69, 118, 167, 235 and 373 sampling sites). Total sampling cost was measured as total time: (number of surveyed sites × site sampling time) + cumulative travel times. Bias was calculated as Renkonen (Percentage) dissimilarity between true and sampled species sets based on area of occupancy (Fig. 2a-c) and species frequency (Fig. 2d-f). Bars represent the 0.025 and 0.975 quantiles: vertical bars refer to distribution of the bias, whereas horizontal bars refer to total sampling times. Blue circles = grid sampling; green squares = random sampling; black diamonds = accessibility-based sampling; grey diamonds = accessibility-based sampling, 600 and 750 sampling sites.

Figure 3 (b/w in print)

Species level: relationships between total sampling costs and bias for the three sampling methods (grid, random and accessibility-based samplings) using the high accessibility scenario. Grid sampling was performed using cell sizes of 10, 6.67, 5, 4, 3.33 and 2.5 km (corresponding to 32, 69, 118, 167, 235 and 373 sampling sites). Total sampling cost was measured as total time (see Fig. 2). Three species were reported: the commonest (Species 1), the rarest (Species 9) and one species with an intermediate frequency (Species 10). Estimates of single species distribution were based on area of occupancy (Fig. 3a-c) and species frequency (Fig. 3d-f). Relative bias was calculated as (true value - estimated value) / true value. Bars represent the 0.025 and 0.975 quantiles: vertical bars refer to distribution of the bias, whereas horizontal bars to total sampling time. Blue circles = grid sampling; green squares = random sampling; black diamonds = accessibility-based sampling; grey diamonds = accessibility-based sampling, 600 and 750 sampling sites.

Figure 4 (b/w in print)

Species level: relationships between total sampling costs and bias for the three sampling methods (grid, random and accessibility-based samplings) using the high accessibility scenario. Grid sampling was performed using cell sizes of 10, 6.67, 5, 4, 3.33 and 2.5 km (corresponding to 32, 69, 118, 167, 235 and 373 sampling sites). Total sampling cost was measured as total time (see Fig. 2). Three species were reported: the commonest (Species 1; Fig. 4a and d), the rarest (Species 9; Fig. 4c and f) and one species with an intermediate frequency (Species 10; Fig. 4b and e). Estimates of single species distribution were based on extent of occurrence (Fig. 4a-c) and niche breadth (Fig. 4d-f). Relative bias was calculated as (true value - estimated value) / true value. Bars represent the 0.025 and 0.975 quantiles: vertical bars refer to distribution of the bias, whereas horizontal bars to total sampling time. Blue circles = grid sampling; green squares = random sampling; black diamonds = accessibility-based sampling; grey diamonds = accessibility-based sampling, 600 and 750 sampling sites.

Tables

Table 1: Ecological preferences and occupancy probability of the 15 artificial species. We assumed Gaussian responses of the species to elevational gradients: each species was characterized by its optimal value (mean) and amplitude of the response (sd - standard deviation). Three species typologies were considered, according to their habitat preferences: specialists for lentic habitats (e.g., ponds), specialists for lotic habitats (e.g., streams) and generalists. Finally, two occupancy probabilities (0.5 and 0.25) were used to control the relative rarity of the species within suitable habitats.

Species	**Elevation range (m a.s.l.)**		**Habitat typology**	**Occupancy**
	mean	**sd**		
Species 1	250	125	lotic	0.5
Species 2	900	200	lotic	0.5
Species 3	1900	300	lotic	0.5
Species 4	600	300	lotic	0.25
Species 5	1150	325	lotic	0.25
Species 6	1600	300	lotic	0.25
Species 7	250	125	lentic	0.5
Species 8	900	200	lentic	0.5
Species 9	1900	300	lentic	0.5
Species 10	600	300	lentic	0.25
Species 11	1150	325	lentic	0.25
Species 12	1600	300	lentic	0.25
Species 13	500	250	lentic + lotic	0.25
Species 14	1000	250	lentic + lotic	0.25
Species 15	1500	250	lentic + lotic	0.25

Electronic Supplementary Materials

Supplementary Data: ESM1 - Data sets and R scripts used to run the analyses

Supplementary Figures: 1-6 (ESM2)