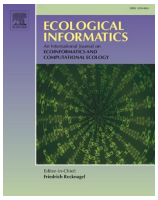
Ecological Informatics 99 (2026) 104112



Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Problems of geometry, sampling, and scale in gridded biodiversity data, and proposed solutions

Petr Keil [a,*], Friederike Wölke [a,b], Gabriel Ortega-Solís [a], Florencia Grattarola [a], François Leroy [a,c], Florian Jansen [d], Carmen Soria [a]

[a] Faculty of Environmental Sciences, Czech University of Life Sciences in Prague, Kamýcká 129, Praha-Suchdol, 165 00, Czech Republic
[b] Department of Ecology, Environment and Plant Sciences, Stockholm University, Svante Arrhenius väg 20 A, Stockholm, 106 91, Sweden
[c] Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH, 43210, USA
[d] Faculty of Agriculture, Civil and Environmental Engineering, University of Rostock, Justus-von-Liebig-Weg 6, Rostock, 18059, Germany

ARTICLE INFO

Keywords:
Biogeography
Scale
Macroecology
Squares
Species-area
bias
Observation error
Imperfect detection

ABSTRACT

Grids, and gridded biodiversity data such as regional or country-level atlases, play a prominent role in ecology, particularly in the study of spatial patterns of species occupancy, geographic ranges, biodiversity, and their drivers and temporal dynamics. However, managing, exploring, and analyzing gridded data presents numerous challenges. Here, we review the problems with gridded data, and the existing solutions. We hope to provide guidance, particularly to early-career ecologists who may otherwise struggle to make sense of the various solutions scattered throughout the literature. We focus on grid-specific problems of sampling (e.g. varying sampling method and effort in space and time, imperfect detection), geometry (e.g. varying grid cell area and shape, positional errors), and scale (e.g. spatial grain and temporal extent). Solutions can be implemented before gridding, during data processing, or during data analysis. These include selecting appropriate projections and grid designs, aggregating or filtering grid cells, and accounting for biases within statistical models. By exposing gaps, trade-offs and single solutions to multiple problems across existing approaches, we further aim to inspire the development of novel modelling frameworks that tackle multiple challenges simultaneously.

1. Introduction

In biogeography, macroecology, and applied ecology, data on species distributions are often arranged in a grid (Fig. 1, Box 1). Prominent examples are gridded atlases (Asher, 2001; Keller et al., 2020; Preston, 2013; Šťastný et al., 2021) (Table 1). Compared to local and spatially discrete data from local observations as in most data in GBIF (www.gbif.org) or in USGS North American Breeding Bird Survey (Ziolkowski et al., 2022), grids have the advantage that their cells are directly adjacent. This enables mapping the data with no blank spots, and it also ensures a thorough, continuous coverage of environments—useful when, for instance, modelling distributions of species that are absent in some suitable locations (Svenning and Skov, 2004), or when assessing occupancy and species ranges (Warren et al., 2001). Further, because grids are contiguous, we can assess not only occupancy, but also spatial patterns such as aggregation or fractality (Gaston, 2003). Importantly, adjacent cells can be aggregated to create coarser cells (Hurlbert and Jetz, 2007; Rahbek, 2005) or can be analyzed using moving windows (Lennon et al., 2001). This is necessary for analyses of scale-dependent phenomena such as species-area relationships (Storch, 2016), effects of abiotic environment on species-environment associations (Belmaker and Jetz, 2011), or occupancy-area relationship (Kunin, 1998), including their change in time (Leroy et al., 2024). Gridded biodiversity data are thus a vital ecological resource.

Managing, exploring, and analyzing data on species distributions in a grid, however, comes with problems. Perhaps the most discussed in the literature is the problem of incomplete or biased sampling (Boyd et al., 2021, 2024; Isaac et al., 2014; Rapacciuolo et al., 2021). Less discussed are problems with geometry (Kunin, 1997) and scale (Boyd et al., 2024; Hurlbert and Jetz, 2007). Furthermore, there may be problems, decisions, and solutions that are not part of methodological overviews, but are described in specific biodiversity analyses in methods sections, appendices, protocols, and workflows. These may be scattered, and motivations for decisions are not always well described or justified: we may see what people did, but we don't know why.

In this paper, we review problems associated with gridded data,

* Corresponding author.
E-mail address: keil@fzp.czu.cz (P. Keil).

https://doi.org/10.1016/j.ecoinf.2026.104112
Received 28 April 2026; Received in revised form 20 October 2026; Accepted 20 October 2026
Available online 21 October 2026
1574-9541/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

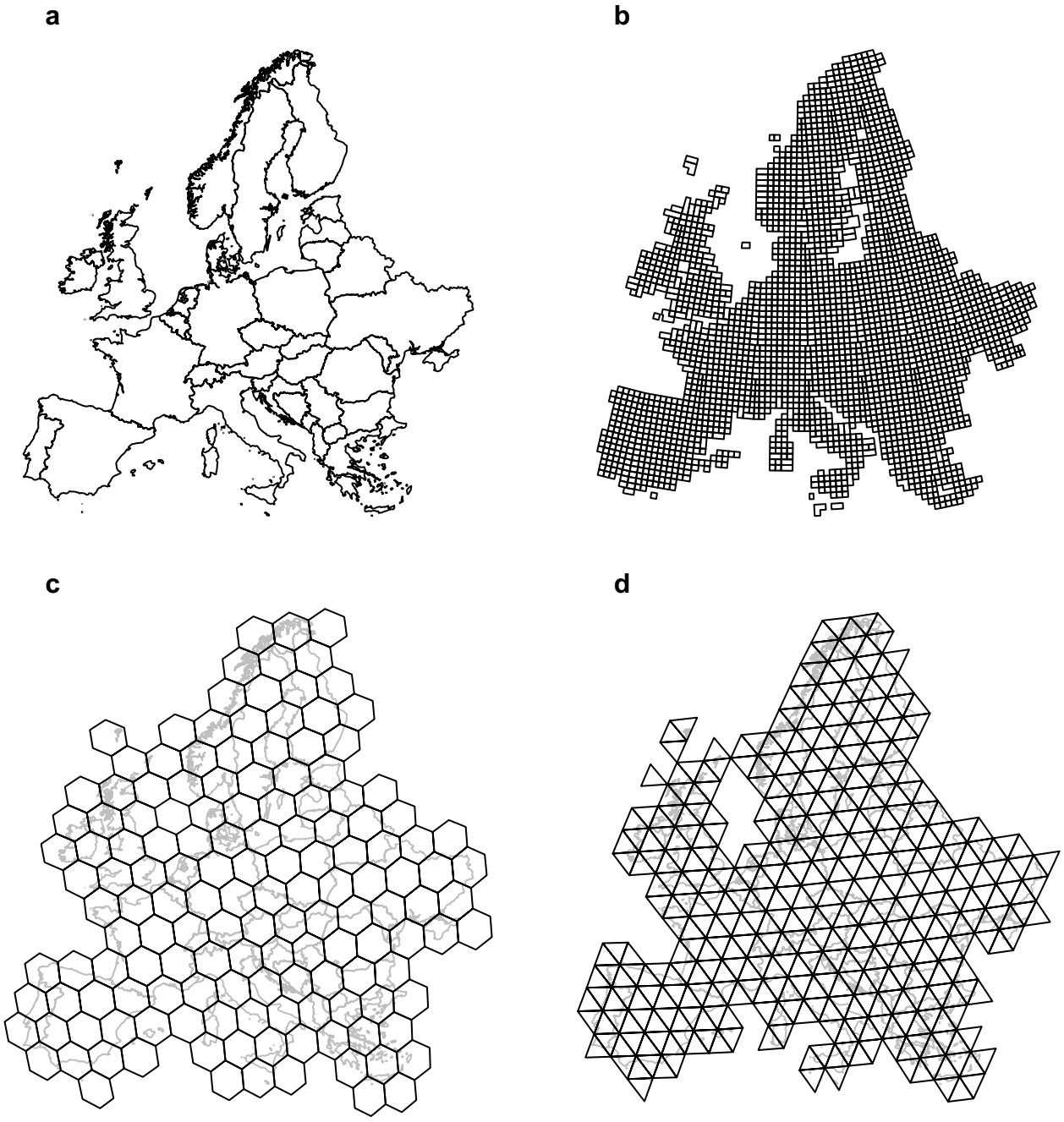


Fig. 1. Examples of spatially contiguous grids. (a) A grid where cells are European countries, and they vary considerably in their area and shape. (b) A rectangular UTM grid used in Keller et al. (2020); note the irregular non-square shapes of some cells. (c, d) Hexagonal and triangular grids; their cells are only minimally distorted by geographic projections. Note that we are not dealing with grids of spatially discrete points in this review.

summarize their causes, and propose solutions. We focus specifically on geometry, sampling, and scale, the areas closest to our expertise. The proposed solutions aim to (i) make raw gridded data more informative in visualizations and exploratory analyses and (ii) make these data more suitable for statistical modelling. Rather than offering a single solution to every problem, we present a toolbox of approaches and show that some address multiple problems simultaneously. By identifying connections among the problems, we expose gaps in current analytical approaches and highlight opportunities for improvement. The problems and solutions apply both to existing grids and gridded atlases and to the design of new grids. Ultimately, our goal is to support accurate maps and adequate models of species distributions and diversity that represent reality and provide reliable predictions and inferences.

Box 1
Basic definitions

Here we use the term **grid** (also called a mesh) for a structure consisting of multiple grid **cells** (polygons) of non-zero area, arranged in a **contiguous** way so that each grid cell is adjacent to another (Fig. 1); we thus leave out point meshes. In our view, grid cells may have varying area and shape even within one grid, and thus we also consider a set of adjacent countries (Essl et al., 2013) or a set of smaller adjacent administrative units (Craven et al., 2020) to be grids.

Gridded biodiversity data describe the distribution of one or more species, or their diversity, across the grid. A typical example is a gridded **atlas** (Table 1) in which data values are always defined per grid cell. These data can be detections or non-detections of species, their presence and/or absences, some metrics or proxies for abundance, species richness, or other indices of diversity.

The *average* area of the cell is the grid's **spatial grain** (resolution), and the distance or area between the grid's outer boundaries is the grid's **spatial extent**. For some ecological quantities (biodiversity and related concepts such as beta diversity, species associations, or species composition), there is no "correct" or "best" grain of a grid, and fine-grain grids are not better than coarse ones, they simply carry different information (Rahbek, 2005). This may come as a surprise to researchers from remote sensing, where the finer the raster resolution, the better. Finally, a **grid origin** is the exact geographic location of a selected fixed point of the grid; if we keep the grain and extent constant, but we change the origin, we are effectively moving the grid around the map.

Structured vs semi-structured vs unstructured gridded data (Table 1). The highest-quality gridded data derive from **structured** surveys with standardized protocols that pre-specify the sampling method, effort, site selection, and temporal replication. **Semi-structured** datasets also derive from surveys that partly standardize sampling protocols and site selection, but they allow some flexibility in the number of observers, the time they spend in the field, the number of visits, and the selection of sites. Importantly, this variation is usually documented in metadata associated with the data collection. Finally, gridded data can be created through **unstructured** sampling processes, such as opportunistic observations, by overlaying point observations on a grid. In this case, the observation process is unknown, and there is little to no metadata on the biases inherent in the data collection. Many museum collections and most citizen science initiatives (e.g., eBird, iNaturalist, Observation.org) gather unstructured point occurrences (Pocock et al., 2017) widely available through GBIF.

2. Literature processing

To identify literature on gridded biodiversity data, we combined the following approaches. (1) In 2023, we searched Web of Science and Scopus for the string "atlas AND biodiversity," which yielded 351 papers. (2) We used the AI tool Elicit to scan these papers and extract information about taxa, temporal replication, data type, geographic extent, spatial scale, and metrics. After reviewing the output, we read all relevant papers individually, examined their cited references, and assigned specific problems and solutions to individual co-authors for further investigation. However, we did not use generative AI to write the manuscript; we wrote it manually. (3) We added 59 papers with which we were already familiar through our extensive experience with the subject, as well as additional relevant papers suggested by colleagues who work with gridded atlases, specifically Nick Isaac, Diana Bowler, and Oliver L. Pescott.

3. Geometrical problems

Problems of geometry affect almost every grid to some degree. These problems include variation of grid cell shape, or variation of the shape of the habitable area within cells (Fig. 2). Many of these are caused by geographic projections. For example, in rectangular grids with global extent and equal-area projections, there is increasing north-south elongation of cells towards the poles. An example is the Behrmann projection of global vertebrate grids (Coelho et al., 2023; Faurby et al., 2018), or the European $50 \times 50\,km^2$ UTM grid (European bird census council, 2020) with increasing fraction of small grid cells northwards. Second, cells may have an irregular shape by definition; these can be countries, biogeographic regions, or counties (e.g. Craven et al., 2020). Finally, problems of geometry often arise when grids are created by aggregation ("coarsening") of other grids, when coarse-grain grid cells are not fully filled by fine-grain cells, particularly around the grid's outer boundary (Keil et al., 2011).

3.1. Overview of geometrical problems

P1: Varying cell area. Some cells in a grid can be larger than others (Fig. 2). This severely affects prevalence (occupancy) of every species, since it follows a positive occupancy-area relationship (Kunin, 1998); in other words, the larger the cell, the larger the probability that a species is in it. Consequently, variation of cell area can mask, or distort, geographic patterns and other drivers of distributions. The same problem applies to species diversity, which follows the species-area relationship (Arrhenius, 1921; Storch, 2016): the larger the grid cell, the more species are in it. Furthermore, variation in cell area distorts beta diversity, species co-occurrences, and species composition in grids, since they are all derived from species distributions and/or from diversity patterns.

P2: Varying cell elongation. Some cells can be elongated more than others (Fig. 2). The more elongated the shape, the larger the variety of environments in the cell, and the more species in the cell. This is because of Tobler's law which states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). The more elongated cells are, the longer gradients of environments and species composition they cover, and species compositional similarity decays with distance (Nekola and White, 1999). This decay is due to varying environmental niches among species, but can also emerge from purely neutral processes in homogeneous environments (Rosindell et al., 2011). Consequently, given constant grid cell area, more elongated grid cells have higher per-species probability of occupancy and more species than square or circular cells (Kunin, 1997), affecting patterns of species occupancy, diversity, and beta diversity. Thus, when elongation of cells varies, it can distort several biodiversity patterns independently of area.

P3: Habitable area is a fraction of cell area. The world is full of boundaries and barriers that delineate habitable areas for species (Fig. 2). An example is the coastline delimiting land for terrestrial species. When we overlay a grid cell over the boundary, we end up with only a fraction of the cell covered by the habitable area; for example, in the grid of the European bird atlas (Keller et al., 2020) the only land in a $\sim 2500\,km^2$ cell is Ibiza, an island of $573\,km^2$, which is 22.9% of the

cell area. Because of this small inhabitable area, the number of terrestrial species in the cell is likely lower than in an inland cell of continental Spain with 100% of mainland. Consequently, even in a regular grid with constant cell area and shape, variation of the fraction of habitable area may distort patterns of species occupancy, distributions, and diversity. Indeed, sharp habitat or biome boundaries may cause the same problem.

P4: Shifted grid origin. When we shift the whole grid (we move the origin of the grid, Fig. 2) by less than a grid cell length, we affect occupancy and apparent distribution of species (Witte et al., 2008), and consequently also patterns of diversity. This problem emerges during creation of new grids, or in already existing older and newer grids in the same region, e.g. in the Czech atlases (Šťastný et al., 2021). The problem of shifted origin gets more serious the coarser the grain of the grid (Witte et al., 2008) because (i) there are fewer cells, and thus smaller sample size for statistical analyses, and (ii) because rare species located close to cell boundaries can completely reverse the spatial pattern.

P5: Small number of grid cells, e.g. at coarse resolutions or small extents. Some grids have a small number of grid cells natively, others because they have been created by aggregation (coarsening) of finer grids (Craven et al., 2020; Rahbek, 2005). If the number of cells becomes too low, it may compromise statistical estimators that need sufficient sample size.

3.2. Solutions to geometrical problems

Some of the solutions (S1–2 and partly S5) are only applicable when we have the luxury of designing the grid, which can either be when a new atlas initiative is launched, or before unstructured point observations (e.g. from GBIF) are aggregated to a grid. Other solutions (S3-S8) are more practical as they also work for already existing grids.

S1: Use equal-area projection. A popular way to avoid varying area caused by geographic projections is to use a regular grid in an equal-area projection; these projections can be custom tailored for any region using projectionwizard.org. However, note that, for rectangular grids, equal-area grids cause variation in elongation, which also affects diversity and distributions (Kunin, 1997); see also P2. This problem grows with increasing geographic extent of the grid, with global grids being most affected. A typical example is the popular Behrman equal-area projection of global grids (Faurby et al., 2018; Jetz et al., 2012), with extremely elongated cells at high latitudes. Thus, for these large spatial extents, hexagonal grids may be preferable (S2).

S2: Use triangular or hexagonal grids. Triangular and hexagonal grids don't require special projections, and they keep area, distance between cell centroids, and shape constant. In other words, they can be wrapped around a sphere while avoiding shape distortion. An example use of a global hexagonal grid is in Keil and Chase (2019). The disadvantage of hexagons is that they look unfamiliar, they cannot be easily converted to rasters to be processed using fast raster-based algorithms, and are challenging to be made nested, i.e. aggregated to coarser resolutions; but see Jurasinski and Beierkuhnlein (2006). Triangular grids can be coarsened and made nested, but we are unaware of any use of a triangular grid in biogeography.

S3: Record the area and shape for each cell. If the previous solutions can't be applied (e.g., because the grid is given and we have no control over its geometry), or when the habitable area is just a fraction of grid cell area, then the geometrical problems should be quantified and noted. We suggest that, for each cell in the grid, metadata are provided that specify: cell area, fraction of habitable area within the cell (e.g. fraction of mainland if terrestrial organisms are mapped), fraction of cell area covered by smaller grid cells (if the grid is an aggregate of a finer grid), and some metrics of shape deformation, e.g. circularity, linearity, or elongation ratio (Stojmenovic and Nayak, 2008). These metrics can then be used as covariates in statistical models (S7), such as in Leroy et al. (2024), or they can be used to filter out problematic cells that don't meet certain criteria (S4).

S4: Remove cells with undesirable geometry. These can be cells with excessive shape distortion (too elongated), cells with a small fraction of land, or cells that are too small or too large. Examples of this are cells overlapping coastal areas or the thin cells—sometimes less than 1 km wide but 10 km long—created when two or more contiguous UTM zones are used to build a grid. A simple and brute-force solution is to remove such cells or merge them with their neighbors, which usually requires a threshold to be specified, for example a threshold of a minimum area of land within a grid cell. Threshold values for this removal can vary in strictness, e.g. only include cells with 30% or with 100% of land (Keil et al., 2012), and the decision about the threshold is often arbitrary or intuitive, and there will likely be a tradeoff between sample size (number of cells) and the threshold strictness. Furthermore, removing cells with small amount of land mass may lead to disproportionate exclusion of coastal grid cells, which may be undesirable when coastal specialist species are of interest, or when the grid covers an island region with complex coastline such as in Japan (Ueta et al., 2023) or New Zealand (Bull et al., 1985) atlases. In these cases, recording the area and shape of landmass and using it as a covariate in models (see S7 below) is preferable to cell removal.

S5: Test sensitivity of patterns to grid origin. This can be wise particularly for coarse grids with few grid cells. When a pattern changes considerably with grid origin, avoid general conclusions from models based on these grids. When a fine-grain grid of the same extent as the coarse-grain grid is available, we can use moving windows (Lennon et al., 2001), i.e. a large sampling cell is moved continuously over the fine grain grid, and values such as diversity are calculated per fine-grain cell in the center of the sampling cell. This is equivalent to overlaying many coarse-grain grids with different origins. However, we should stress that biodiversity patterns specific to a particular grid origin are not incorrect, even when different from patterns observed in other grid origins. A particular grid origin may even be practical and preferable, e.g. when it leads to more even geometry of habitable area within cells, or when it matches a certain administrative division of a region.

S6: Coarsen up the grid. Some problems of varying cell area and shape can be fixed by aggregation of smaller cells to larger units, as done for example by Craven et al. (2020). This can be also done to ameliorate problems of positional errors (Moudrý and Šímová, 2012), although there is some controversy around that (Gábor et al., 2022). The universal downside of coarsening up the grid (see also S14 and S21) is reducing geographical detail and sample size; however, sometimes this may be a price worth paying for the increased reliability of diversity and distributions at coarse resolutions.

Table 1
Examples of gridded atlas datasets, ordered from most structured (top rows) to least structured (bottom rows). Note that the Japanese bird atlas also includes additional unstructured data, which can be, however, filtered out.

Dataset	Type	Spatial extent	Spatial grain	Temporal extent	Sampling effort	Effort units per cell
Birds of Japan (Ueta et al., 2023)	Structured	Japan	$20 \times 20\,km^2$	2016–2021	Known, fixed	
Czech bird atlas (Šťastný et al., 2021)	Semi-structured	Czech Republic	$11.2 \times 12\,km^2$	2014–2017	Known but varies	# cards, # observers
SABAP2 - South African bird atlas (Brooks et al., 2022)	Semi-structured	South African countries	5′ × 5’	2007-present	Known but varies	# cards, hours
Mammals of Hamburg (Schäfers et al., 2016)	Unstructured	Hamburg, Germany	$2 \times 2\,km^2$	1997–2015	No	

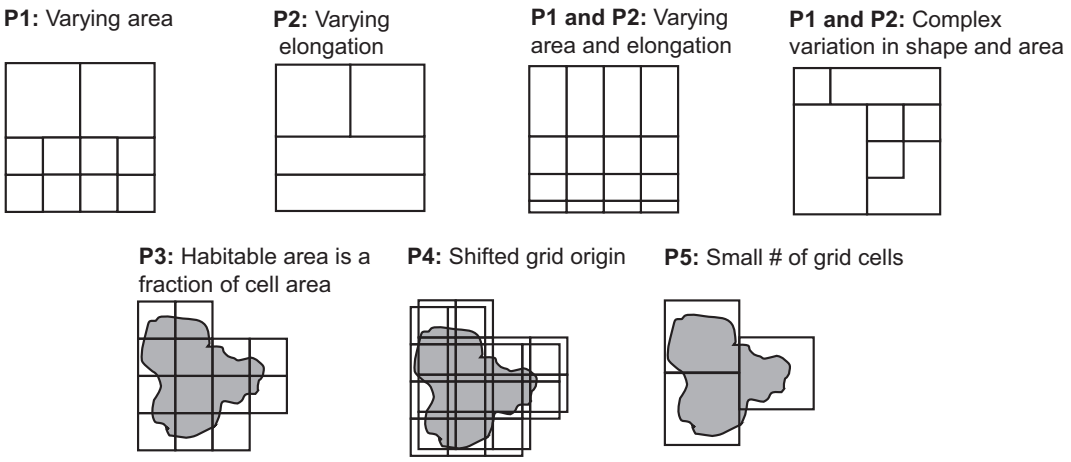


Fig. 2. Common geometrical problems of gridded biodiversity data.

S7: Use area and shape in statistical models. When cells vary in their *area* (both, the cells themselves, and the proportion of the cell covered by habitable land), there are ways to include the effect of area in statistical models. First, a mathematical transformation of area values may be necessary and we strongly recommend it, usually log- or square root transformations. Second, the transformed area can be used as a predictor in models of diversity and distributions, ideally also as an interaction term with other predictors, e.g. see Keil and Chase (2019) and also S24. The latter stems from the well-established observation that sample area not only affects diversity and distributions per se (Storch et al., 2007), but it also modulates the effects of other variables (Belmaker and Jetz, 2011). Furthermore, some types of hierarchical statistical models assume a very specific effect of area so that it is used in models as an offset (Grattarola et al., 2023, 2024; Isaac et al., 2019). Regarding the *shape* attributes, a common problem is varying elongation, which is common in both rectangular and irregular grids. Compared to area, there is much less research on how (and if) elongation should be transformed, and how it interacts with other predictors. Here, we recommend standard model selection procedures to assess the effect of elongation, as well as investigating the statistical interaction between shape and area.

S8: Avoid generalizations from a few cells. When the number of cells is small, for example at very large grains, it may still be useful to calculate some quantities per grid cell; for example, we may be interested in the total number of species (species pool) in a large region corresponding to a large grid cell. However, statistical generalizations based on these small samples are not advisable. If generality is desirable, use a finer grid with more grid cells. Alternatively, Bayesian statistics can quantify the uncertainty caused by small sample size (Gelman et al., 2013).

4. Sampling problems

Sampling problems (Fig. 3), specifically those related to variation in sampling effort and methods, are perhaps the most common and influential problems in gridded biodiversity data. They were also a subject of several reviews (Botts et al., 2012; Boyd et al., 2023; Isaac et al., 2014; Robertson et al., 2010). The ultimate reason for sampling problems is that we mostly cannot survey an entire grid cell, nor can we survey it continuously over time. Instead, we are left with spatially and temporally discrete samples (plots, point counts, surveys, visits) or incidental records. The total number of such samples per grid cell, combined with the total number of individuals per sample, sample area or length (when the sample is a transect), or time spent sampling, represent *sampling effort*, and the lack of information on sampling effort is behind most of the problems listed in this section.

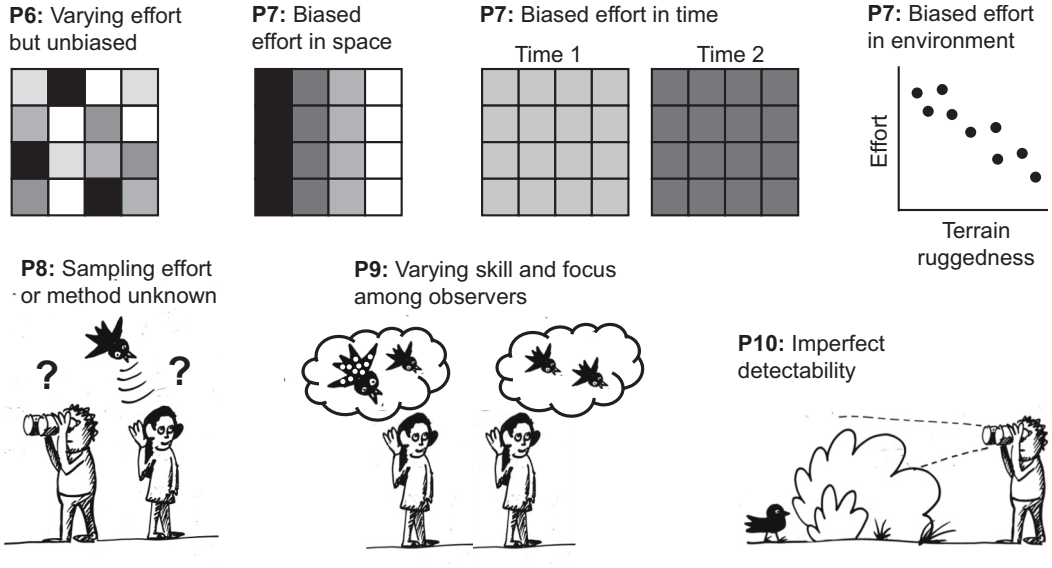


Fig. 3. Common sampling problems of gridded biodiversity data.

4.1. Overview of sampling problems

P6: Varying sampling effort. Sampling effort can vary geographically, temporally, or spatiotemporally, which can render some grid cells in certain periods under-sampled. This may affect some ecological quantities, e.g. occupancy of species, average richness, or regional richness. However, if the variation is random in space (Fig. 3), time, or environment, it may not be such a problem for geographic, temporal, or environmental gradients of diversity and distributions. However, this randomness will likely be rare.

P7: Biased sampling effort in space, time, or environment. This is a special case of the abovementioned problem, when sampling not only varies, but is also systematically biased (clumped) in space (Fig. 3), time (Fig. 3), or both, for instance when some regions are better accessible or attractive. Temporal biases (Bowler et al., 2022) (Fig. 3) can emerge when the weather (or climate) in certain time periods is better for sampling, or with the general increase in biodiversity records towards the present (https://www.gbif.org). This issue might also arise due to a shift in sampling focus from one area to another, e.g., when one region is already well-sampled, and the surveyors move to another region. Even worse, there can be a joint bias both in time and space, with some regions sampled only in a certain period. This can be coupled with biases in environmental space (Xu et al., 2024), e.g. when some environments are less accessible or attractive to observers (Fig. 3). All these biases lead to detected patterns of diversity and distributions that differ from real-world patterns. In the most extreme, real-world patterns are then completely obscured or reversed by patterns of sampling (Grattarola et al., 2020).

P8: Sampling effort or method unknown. Sometimes we may have no information on the sampling effort. This can be further complicated when sampling methods are unknown, or insufficiently described. In such cases it may seem impossible to determine if non-detections of species are real absences (Dobson et al., 2020; Rapacciuolo et al., 2021).

P9: Varying skill and focus among observers. Ecological data collection usually involves many observers with various levels of expertise, impacting data quality (Isaac et al., 2014). Even when considering systematic data collection, less-skilled observers may work slowly, are more likely to miss cryptic species, and tend to mis-identify species. Consequently, even if sampling effort in observer hours per grid cell is constant, the real sampling may vary, i.e. it is lower in cells surveyed by unskilled observers (Prendergast et al., 1993). Since identification relies on visual and auditory cues, the ability of observers to detect the species around has been shown to decay with age (Farmer et al., 2014). Apart from skill, the focus may vary between observers, some preferring species from a certain habitat or certain type of taxa, e.g. some observers may be biased against alien species, or towards conspicuous taxa (Bowler et al., 2022; Johnston et al., 2023).

P10: Imperfect detectability. Detecting all individuals in a grid cell is not guaranteed due to combination of imperfect detectability and insufficient sampling effort, resulting in false negatives and inaccurate distributions (MacKenzie et al., 2002). The probability of detection at a grid cell depends on species abundance, conspicuity, and behavior, which are influenced by the cell characteristics (habitat, disturbance level), environmental conditions (including season, phenology, weather, time of the day), and survey method, timing, effort, and observer skill (Guillera-Arroita, 2017). As perfect surveys are unlikely, most raw gridded biodiversity data tend to have a lower number of detected presences than true occurrences of species.

4.2. Solutions to sampling problems

Solutions to sampling problems nearly always involve some estimation of sampling effort as the first step (S9-S12) and vary from adjustments of the data such as exclusion of under-sampled grid cells (S13) or coarsening up the grid (S14), to custom-tailored algorithms (S15-S16), and model-based solutions based on hierarchical models (S17-S20). The latter is both flexible and mathematically sound, but hierarchical models can also be challenging to specify correctly, can have high computational cost, and can come with steep learning curve. Thus, we encourage researchers to also consider simpler or more ad-hoc solutions, and/or to creatively combine them with hierarchical models.

S9: Get data on sampling methods and effort. We advise seeking data on sampling effort and methods in gridded biodiversity data whenever possible; if unavailable publicly, they may still be possible to get from the authors. Data on effort can take many forms, e.g. number of traps, time of trap exposure, number of transects and their length, or surveyed point locations. In semi-structured atlases (e.g. the Czech bird atlas; Ortega-Solís et al., 2026), proxies of effort can be the number of survey cards per cell (cards containing lists of species detected during a visit), number of observers per grid cell, or total number of observer hours over all visits in a cell. Sampling effort may be further standardized to a relative measure of sampling *completeness* (Keller et al., 2020). These metrics of effort can be then used in statistical models as covariates, as we describe below. Here we also caution that even though some of these data may be close to a real effort, they may not be the exact measure of it. For instance, just the number of observers per grid cell is not capturing the variation in hours per observer. Yet, even limited information on the reporting rate can improve models (Isaac et al., 2014).

S10: Get data on recorder identity. Knowing the identity of recorders can be used to determine *recorder syndromes* (Isaac and Pocock, 2015), i.e., typical sets of recording behaviors which can be estimated from species list lengths, coverage of rare or difficult-to-identify species, frequency of recording etc. These can then be used in statistical models that account for sampling variation (Isaac and Pocock, 2015) (S17, 18). However, recorder identity is sensitive personal information that is not usually freely available, not even in rare cases when it is known. Providing anonymized recorder IDs (e.g. recorder 547) could solve this issue.

S11: Use external proxies of sampling effort. In unstructured or semi-structured data, sampling effort often correlates with accessibility, e.g. modelled as distance to human settlements (Monsarrat et al., 2019) or distance to roads and research institutions (Meyer et al., 2015). Not commonly used but potentially useful may be the distance from attractive protected areas where observers enjoy working more than in cultural landscapes. These data can be extracted from public GIS layers and used as proxies for sampling effort in hierarchical models (S17, 18). The downside is that these proxies alone can only inform about relative effort values (e.g., better accessible grid cells are more likely to contain detections), but not about the absolute effort which would enable probabilistic statements about the actual probability of presence of the species.

S12: Derive proxies of effort from detections. Some methods use detections and non-detections of species to infer proxies of sampling effort. For instance, Szabo et al. (2010) proposed to use the number of species detected in a grid cell as an indicator of sampling effort; we see this method as questionable, since it confuses effort with one of the fundamental ecological quantities of interest. In a better alternative, Hill (2012) proposed to use the frequency of the so called “benchmark species” that are detected in the neighboring cells as an indicator of sampling effort. Encouragingly, we found that this frequency correlates well with some real metrics of sampling effort (under review). Hill (2012) then demonstrates how this can be used, in what is now known as the *Frescalo algorithm*, to infer species occupancy trends over time or to identify unexpected presences or absences. Studies that use the algorithm are Eichenberg et al. (2021) or Montràs-Janer et al. (2024). There is also a practical guide on how to use the algorithm (Goury et al., 2026). This proxy of effort can then be used in hierarchical models (S17-S20), although the same limitation applies as for the external proxies mentioned above: the Frescalo-derived proportion of benchmark species is a proxy of relative sampling effort, not absolute one (see S11).

S13: Remove (filter out) under-sampled cells. Once we have the

sampling effort (S9), or its relative proxy (S11, 12), a crude but simple solution is to remove sites with sampling effort below a certain threshold. One way to determine the threshold is to model species richness as a function of the effort, and to determine where the function approximately flattens out (Prajzlerová et al., 2024). Indeed, removing cells means considerable data loss, particularly when assessing temporal change across multiple time periods (atlas editions), since even well-sampled cells in one atlas period may need to be removed if the same cells were under-sampled in another period. A good situation for this solution is when the total number of grid cells is high, e.g. at fine grains, and thus even exclusion of large number of under-sampled grid cells retains a large and representative sample of well-sampled cells. A good overview of filtering approaches is provided by Hughes et al. (2024).

S14: Coarsen up the grid. In macroecology, it has been recognized (Hurlbert and Jetz, 2007) that imperfectly sampled fine-scale data converge to true presences and absences if they are aggregated to sufficiently coarse grids. Coarser grids also give more accurate estimates of species occupancy, i.e. the proportion of occupied grid cells (Boyd et al., 2024). In other words, there is a minimum reliable grid resolution (Keil et al., 2014) above which we are safe to assume true presence and absence. This has been demonstrated to work for spatially random variation in sampling (Fig. 3) (Fig. S1 in Keil et al., 2014). This is also the reason why global biodiversity grids have grains of $100 \times 100\ km^2$ and coarser (Faurby et al., 2018; Orme et al., 2006) and why finer global grids (Jenkins et al., 2013) are problematic. As we also mention in S6, S8, S14 and S21, the downside of the increased reliability of coarser grids is the loss of detail and sample size.

S15: Use rarefaction. In a set of studies related through their data and authors (Biesmeijer et al., 2006; Carvalheiro et al., 2013; Keil et al., 2011), a variant of *individual-based rarefaction* (Gotelli and Colwell, 2001; Hsieh et al., 2016) was used to account for variation in sampling effort in gridded data, where records in each grid cell were down-sampled to a standard number to obtain standardized number of species per grid cell. Another example of the use of individual-based rarefaction over a country-wide grid is in Grattarola et al. (2020) or Klímová et al. (2025). It should also be possible to perform *sample-based rarefaction* (Gotelli and Colwell, 2001) to standardize effort across grid cells when biodiversity data are not just individual detections, but sampling events with lists of species per event (e.g. bird point count data or botanical plots). However, beware that we still lack a comprehensive validation of the rarefaction approach in the gridded large-scale setting either on simulated data, or on data with known true richness. Furthermore, the use of rarefaction is only justifiable when detections and/or samples within grid cells are not spatially too clumped, and/or distributed in an unrepresentative subset of environments in the grid cell.

S16: Ad-hoc corrections of metrics of change. Isaac et al. (2014) provide an overview of simple statistical tricks that have been proposed to control for varying effort directly in the calculation of metrics of temporal change. Examples of these include Telfer et al. (2002)'s estimation of change in species occupancy as a deviation from the overall occupancy trend of all species, or Ball et al. (2011)'s proposition to model species status in a cell as the proportion of total records in the grid cell. Although inspirational and potentially easy to implement, these are specific solutions linked to specific metrics and data types and should be used with caution.

S17: Use sampling effort as a covariate. An easy-to-apply method of controlling for sampling problems is to use additional data on sampling effort, method, or their proxies, in simple statistical models such as GLM or in machine learning. The formula of this approach is $Y = f(X, effort)$, where Y is the ecological quantity of interest (e.g. occupancy or species richness) and X are predictors not related to sampling. This approach is valid, but it has a key limitation: It does not estimate the true occurrence of species (or species richness); it only estimates frequency of detections and richness of detected species, given a fixed value of sampling effort. We thus need to assume that frequency of detections corresponds to the actual species occurrence, which may not hold (MacKenzie et al., 2002). Further, the effect of effort on ecological quantities can be non-linear (Gotelli and Colwell, 2001) and effort also interacts in its effect with other variables, sometimes those affecting species detection (Kéry and Royle, 2015), making such simple statistical approach problematic.

S18: Use hierarchical imperfect detection models. There is a class of hierarchical models (Guillera-Arroita, 2017; MacKenzie et al., 2002) which separate the observation (detection) process from the true but unobserved ecological quantity such as occurrence or abundance of a species. These models can be used to account for varying effort and imperfect detection when estimating the probability of a species' presence. These models typically rely on repeated site visits, but these are not always available. One solution is to create "pseudo" visits by dividing unstructured or semi-structured records within an atlas period to shorter time intervals (Klinkovská et al., 2024), but for this we need to know the dates of the records. Another option is to use single-visit versions of occupancy models (Lele et al., 2012) which have been already applied to gridded atlas data (Peach et al., 2017; Sadoti et al., 2013); they either rely on external predictor of the observation process (it needs to be orthogonal to predictors of the ecological process), or on space-for-time substitution where adjacent grid cells are aggregated, which reduces the effective sampling size. Importantly, Pescott (2026) presents the connection between imperfect detection models and the Frescalo algorithm of Hill (2012), see also S12. The practical downside of hierarchical imperfect detection models used to be their steep learning curve and computational cost. However, this has become less of an issue lately, with convenient software packages such as spOccupancy (Doser et al., 2022) and spAbundance (Doser et al., 2024).

S19: Leverage external data with integrated models. Sometimes there can be other data types available for the region of the gridded atlas, and these can improve models that are based on the atlas data with varying sampling effort. Specifically, there are species distribution models which can calculate joint likelihood of multiple types of biodiversity data (Dorazio, 2014; Isaac et al., 2019), and we call these the *integrated models*. These have been used to integrate gridded records with local surveys to improve predictions of population trends (Pagel et al., 2014) or to integrate gridded records with expert range maps and camera traps to predict large-scale dynamics of geographic ranges (Grattarola et al., 2024). For an overview of the uses of integrated models in gridded citizen science data see Johnston et al. (2023). A convenient software package for integrated models is spOccupancy (Doser et al., 2022), although more complex data structures may require custom-coded implementations (Kéry and Royle, 2016, 2020).

S20: Borrow strength from other taxa in joint models. Joint species distribution models [JSDM; (Ovaskainen et al., 2017; Pollock et al., 2014)] are another tool that can strengthen the inference and predictions of true species distributions from data affected by the sampling issues. In the case of gridded biodiversity data affected by sampling issues, the idea is that different species may be affected by the same sampling issues, and thus data on other species can improve estimation of the sampling (observation) process (Rapacciuolo et al., 2021). This is closely related to the abovementioned Hill's (2012) Frescalo algorithm, which also assumes a sampling process shared across species. Furthermore, thanks to niche conservatism, phylogenetically or functionally similar species may also have similar environmental requirements; modelling this in JSDMs can improve estimation of the ecological process (McInerny and Purves, 2011; Ovaskainen et al., 2017) even if the target species is undersampled. A related idea is to model ecological quantities at multiple taxonomic levels (Adjei et al., 2024). Examples of software packages available for JSDMs at large scales are HMSC (Tikhonov et al., 2020), sjSDM (Pichler and Hartig, 2021), or spOccupancy and spAbundance mentioned in S18 and S19 (Doser et al., 2022, 2024).

5. Problems of scale

Scale has several meanings. Here, we deal with two aspects of scale: (1) spatial and temporal *grain* or resolution, which is an average area of a grid cell or average temporal duration, also known as time span (Mellado-Mansilla et al., 2026), of an observation, and (2) spatial and temporal *extent*, which is the total area of a region covered by a gridded dataset, or the total temporal period covered by a series of gridded atlases in the same region. For instance, the last three editions of Czech bird atlases (Ortega-Solís et al., 2026; Šťastný et al., 2021) has spatial grain of 11.2×12 km, spatial extent is the whole Czech Republic, temporal grain is roughly 4 years (the average time covered by one edition), and a temporal extent of 30 years (from 1985 to 2015).

Most scaling problems emerge when comparing gridded datasets, or their editions, with different grains, or when generalizing results from a gridded dataset outside of its extent and grain. The problem of varying (or limited) extent is a problem of generality: the larger the spatial extent of a grid, the more general the pattern (or predictions, or inference) derived from that grid (Wiens, 1989). Hence, problems of extent are mostly solved by not generalizing (extrapolating) outside of the extent of the data. More fundamental are problems of grain (Fig. 4), since different grains reflect different ecological mechanisms. For this reason, we focus on problems and solutions related to spatial and temporal grain, not extent.

5.1. Overview of problems of scale

P11: Varying spatial grain in space or time. Spatial grain of the grid affects species commonness and rarity (Hartley and Kunin, 2003), patterns of species diversity (Rahbek, 2005), turnover (Keil et al., 2012), aggregation (Soria et al., 2026), and also associations of these with the environment (Belmaker and Jetz, 2011). This also includes species-distribution models (SDM), which can lead to different inference and predictions depending on the grain of the data (Mertes and Jetz, 2018). The effect of temporal grain on patterns of species diversity and distribution is much less studied, but there are good theoretical and empirical reasons for temporal grain being similarly important as spatial grain (Mellado-Mansilla et al., 2026), mostly due to the species-time-area relationship (Adler et al., 2005). All this means that analyses of gridded data at one grain may be hard to compare to analyses using different grains.

P12: Collinearity between spatial grain and time span. We expect that the larger the spatial grain of gridded data, the longer the time span (Mellado-Mansilla et al., 2026) over which the data were aggregated (Fig. 4). Examples are European plant grids: The per-country data on European plant colonizations and extinctions (Essl et al., 2013) cover a broad time period of around 100 years, similarly to the 50×50 km European gridded plant atlas (Lahti and Lampinen, 1999) which covers a long period of time of more than 100 years, while smaller per-country gridded data [e.g. Stroh et al., 2023b or Klinkovská et al., 2024] cover intervals of decades. Although fine-grain data can be possible to aggregate to longer time periods, as far as we are aware this has not been done; conversely, we have not heard of a per-country system of regional species lists (as in Fig. 1a) that would cover only a short period of a year or two. When comparing analyses based on data with different spatial grains, one may then observe differences that may in fact be caused by different temporal spans.

P13: Unclear temporal span. Some gridded data are aggregations of all historical records that are available, with unspecified start year of the aggregation (Fig. 4). Examples are the per-country butterfly and plant species lists (Essl et al., 2013; Van Swaay and Warren, 1999) or the abovementioned European plant atlas (Stroh et al., 2023a). This can pose a challenge when assessing rates of change between two periods where the first period is undefined (Keil et al., 2018).

P14: Varying temporal span. Sometimes multiple editions of an atlas in a region can vary in the temporal extent that they cover. For instance, the four editions of the Czech bird atlas (Ortega-Solís et al., 2026; Šťastný et al., 2021) cover 5, 5, 3, and 4 years respectively. If used for assessments of temporal trends, a declining trend may be observed over the four editions simply because the first editions have longer time span.

5.2. Solutions to problems of scale

Below we offer two types of solutions to the problems of scale. The first ones (S21-S22) are mechanical adjustments of the grid aiming at making grains comparable, they are simple but may reduce sample size and fine-grain detail. The second type (S22-S25) are model based solutions which enable inference and predictions at different grains, even at grains not provided in the data, but can be challenging to implement.

S21: Coarsen up the grid. The simplest solution to the variation in spatial grain is to aggregate some or all grid cells to make them comparable with other coarse-grain data. This may also fix some of the geometrical and sampling issues (as described in the previous sections). Apart from the reduced sample size already mentioned in S6 and S14, the downside is that analyses of coarsened data will capture ecological

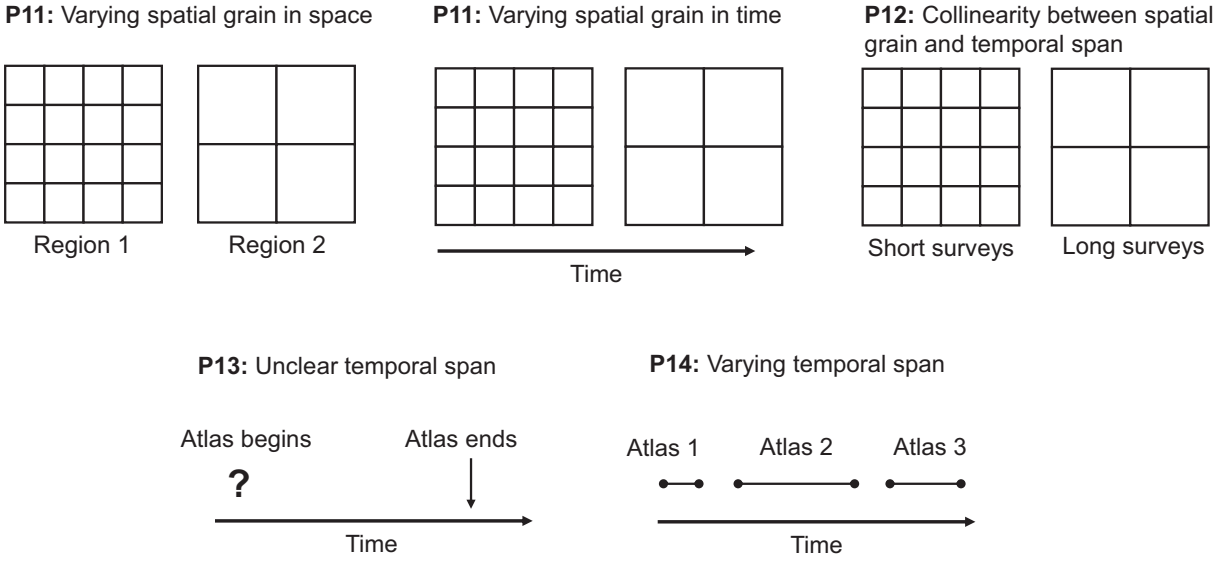


Fig. 4. Common problems of gridded biodiversity data related to spatial and temporal scale.

processes that may differ from fine-grain processes. That does not mean they are incorrect, just different. For instance, species diversity and distributions at fine grains may be reflecting local stochasticity, while at coarse grains they will follow climatic gradients (Craven et al., 2020). Thus, predictions and inference based on models fitted to the coarsened data are not directly comparable with those from fine-grain original data.

S22: Extend time span. An analogical solution to the coarsening up the grid in space is merging multiple temporal replications to one, i.e. extending the time span. This can fix the issue of varying time span (Fig. 4), but comes at a cost of losing temporal detail, potentially also temporal replication, and it is only possible when many temporal replications are available, which can be a rare commodity. To see how extending the time span affects patterns and drivers of biodiversity see Mellado-Mansilla et al. (2026).

S23: Downscale models. A more advanced and challenging

	Geometrical problems	P1: Varying cell area	P2: Varying cell elongation	P3: Habitable area a fraction of cell area	P4: Shifted grid origin	P5: Small # of grid cells	**Sampling problems**	P6: Varying sampling effort	P7: Biased sampling in space, time, environment	P8: Sampling effort or method unknown	P9: Varying skill and focus among observers	P10: Imperfect detectability	**Problems of scale**	P11: Varying spatial grain in space or time	P12: Collinearity of spatial grain and temp. extent	P13: Unclear temporal span	P14: Varying temporal span
Solutions																	
S1: Use equal area projection		●															
S2: Use triangular or hexagonal grid		●	●			●											
S3: Record area and shape of cells		●	●	●										●			
S4: Remove cells with undesirable geometry		●	●	●		●											
S5: Test sensitivity of patterns to grid origin					●												
S6, S14, S21: Coarsen up the grid		●	●	●		●		●	●	●	●	●		●			
S7: Use area and shape in stat. models		●	●	●										●			
S8: Avoid generalizations from a few cells					●	●											
S9: Get data on sampling effort and methods								●	●	●	●	●					
S10: Get data on recorder identity											●						
S11: Use external proxies of sampling effort								●	●								
S12: Derive proxies of effort from detections								●	●	●							
S13: Remove (filter out) undersampled cells								●	●								
S15: Use rarefaction								●	●	●							
S16: Ad-hoc corrections of metrics of change								●	●	●							
S17: Use effort as a covariate								●	●	●	●	●					
S18: Use hierarchical imperfect detection models								●	●	●	●	●					
S19: Leverage external data with integrated models								●	●	●	●	●					
S20: Borrow strength from other taxa in joint models								●	●	●	●	●					
S22: Extend time span								●	●	●	●	●				●	●
S23: Downscale models		●	●											●			
S24: Use spatial grain or time span as a covariate														●			●
S25: Estimate scaling relationships		●	●	●		●								●			

Fig. 5. Connections between the problems and solutions. Black cells indicate that a solution (row) can help with a given problem (column). Red cells indicate that the solution would make the problem worse. The collinearity between spatial and temporal grain is the only problem for which we do not see a solution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

approach is to infer fine-grain distributions or diversity, or ecological processes, from the coarse grain data. This is typically done through hierarchical statistical models (see also S18) where likelihood of fine-grain distributions and parameters are conditional on coarse-grain data (Keil et al., 2013; Keil and Jetz, 2013), or when we assume a scale-free underlying point pattern intensity behind the grid (Isaac et al., 2019). The latter approach also allows integration of gridded data with other types of data for improved inference and predictions (Grattarola et al., 2023; Pagel et al., 2014). While conceptually elegant, the challenge of downscaling models is their computational cost and steep learning curve. On the other hand, if they work, these models make inference and predictions possible even at grains not represented by the data.

S24: Use spatial grain or temporal span as a covariate. When combining entire datasets (atlases) with different grid resolutions in a single analysis, we can include the grain as a covariate, or as an interaction term in statistical models, and we can then fix the grain to a constant when making inference or predictions (Keil and Chase, 2019; Midolo et al., 2025). In the temporal context, this has been done by Mellado-Mansilla et al. (2026). However, using grain as a covariate is only sensible if differences in grain between atlases are not collinear with other properties (such as geographic or environmental gradients). This solution has lower computational costs and is more user friendly than S23 proposed above.

S25: Estimate scaling relationships, not single-scale effects. Rather than a specific solution, this is more of a mindset to approaching biodiversity data, an approach where we give up the idea that there is a single "best" or "correct" relationship or mechanism behind biodiversity. Instead, relationships and mechanisms change continuously as we move from fine to coarse grains. We know that quantities such as species richness (Storch, 2016), their patterns such as latitudinal or altitudinal gradients (Rahbek, 2005), and their environmental drivers change with spatial grain (Belmaker and Jetz, 2011; Field et al., 2009; Keil and Chase, 2019). Let's call this relationship between grain and ecological quantity, or model coefficient, a scaling relationship. These scaling relationships put both the statistical inference and predictions in context of a given grain, enabling comparisons across datasets of different grains. Here, gridded biodiversity data are especially suitable for estimation of these scaling relationships, as they are spatially contiguous and can be coarsened to a set of nested scales. For examples of this practice see Hartley and Kunin (2003), Hurlbert and Jetz (2007) or Keil et al. (2013).

6. Perspectives and future directions

Fig. 5 provides a synthesis of the relationships between the problems and solutions reviewed in this paper. An important insight emerging from this synthesis is that the connection between problems and solutions is many-to-many: most challenges can be addressed by multiple approaches (especially challenges arising from sampling effort), and several approaches that we mention mitigate multiple challenges simultaneously, prominent examples being getting data on sampling effort or coarsening the grid. The latter ameliorates several sampling and scaling problems; however, here we also need caution as sample size may drop too much at too coarse resolutions. Nevertheless, this highlights opportunities to move beyond the individual treatment of geometry, sampling, and scale, towards a solutions for challenges from several domains at once. Hierarchical, integrated, joint, and cross-scale modelling approaches provide promising examples of this direction, but substantial opportunities remain for their development, particularly in terms of improving computing efficiency and user friendliness.

We also note that we have focused on problems that we know from our work, some of which are notorious (such as the sampling problems), whereas other problems only received limited attention (such as the problems of geometry and scale). However, there are whole families of important issues that we have omitted. Examples are problems of taxonomic standardization and taxonomic scale, problems of data ownership and licensing, or large-scale biases in geography and socio-economic history (e.g., most gridded data come from wealthy countries). Each of these would indeed deserve their own section, or even a paper.

7. Conclusion

This review demonstrates that the challenges associated with gridded biodiversity data are diverse but often manageable. Among the solutions discussed, several stand out because they are relatively straightforward to implement or can mitigate multiple problems at a time. These include: (1) Quantifying grid geometry (cell area, shape, and land proportion), (2) characterizing sampling effort through direct measurements or proxies (e.g. using the Frescalo algorithm of Hill (2012), or external variables such as accessibility or distance to institutions/cities), (3) choosing an appropriate spatial grain while considering the tradeoff between sample size and bias correction, and (4) applying modelling approaches that can account for geometry, sampling and scale, for example hierarchical occupancy models, integrated models, joint species distribution models, and cross-scale models.

At the same time, some challenges may be inherently difficult to overcome or have no solution at all. There are collinearities of all sorts (Dormann et al., 2013) and a lack of information about spatial or taxonomic variation in methods or abrupt changes in methodology between editions of atlases. In such cases, we can only hope to get some additional data or information. Furthermore, every solution reviewed here comes with assumptions and limitations, and hidden sources of variation and bias will inevitably remain.

Nevertheless, gridded biodiversity data remain among the most valuable resources available for understanding species distributions, diversity patterns, and their dynamics. By recognizing their limitations and applying appropriate solutions, researchers can substantially improve ecological inference and generate more reliable maps, predictions, and conservation insights.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the author(s) used AI tool Elicit to look for relevant papers (see the Literature processing section above). We did not use generative AI to write the manuscript; we wrote it manually. We also created all the figures manually.

CRediT authorship contribution statement

Petr Keil: Writing – review & editing, Writing – original draft, Conceptualization. **Friederike Wölke:** Writing – review & editing, Writing – original draft, Conceptualization. **Gabriel Ortega-Solís:** Writing – review & editing, Writing – original draft, Conceptualization. **Florencia Grattarola:** Writing – review & editing, Writing – original draft, Conceptualization. **François Leroy:** Writing – review & editing, Writing – original draft, Conceptualization. **Florian Jansen:** Writing – review & editing. **Carmen Soria:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Petr Keil reports financial support was provided by European Research Council. Friederike Wölke reports financial support was provided by European Research Council. Gabriel Ortega reports financial support was provided by European Research Council. Florencia Grattarola reports financial support was provided by European Research Council. Francois Leroy reports financial support was provided by European Research Council. Carmen Soria reports financial support was provided

by European Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

P. K., F. W., G. O., F. G., F. L. and C. S. were supported by the European Union (ERC), BEAST, 101044740. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the Eusropean Union nor the granting authority can be held responsible for them. We thank Oliver L. Pescott and Melanie Tietje for helpful comments.

Data availability

No data or code were used in this article.

References

Adjei, K.P., Carvell, C., Isaac, N.J.B., Mancini, F., O'Hara, R.B., 2024. Integrating data from different taxonomic resolutions to better estimate community alpha diversity. Ecography 2024, 1–13 e07182.
Adler, P.B., White, E.P., Lauenroth, W.K., Kaufman, D.M., Rassweiler, A., Rusak, J.A., 2005. Evidence for a general species–time–area relationship. Ecology 86, 2032–2039.
Arrhenius, O., 1921. Species and area. J. Ecol. 9, 95–99.
Asher, J. (Ed.), 2001. The Millennium Atlas of Butterflies in Britain and Ireland. Oxford University Press, New York.
Ball, S.G., Morris, R.K.A., Rotheray, G.E., Watt, K., 2011. Atlas of the hoverflies of Great Britain (diptera, syrphidae), Centre for Ecology & Hydrology.
Belmaker, J., Jetz, W., 2011. Cross-scale variation in species richness–environment associations. Glob. Ecol. Biogeogr. 20, 464–474.
Biesmeijer, J.C., Roberts, S.P.M., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., Schaffers, A.P., Potts, S.G., Kleukers, R., Thomas, C.D., Settele, J., Kunin, W.E., 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science 313, 351–354.
Botts, E.A., Erasmus, B.F.N., Alexander, G.J., 2012. Methods to detect species range size change from biological atlas data: A comparison using the south African frog atlas project. Biol. Conserv. 146, 72–80.
Bowler, D.E., Callaghan, C.T., Bhandari, N., Henle, K., Benjamin Barth, M., Koppitz, C., Klenke, R., Winter, M., Jansen, F., Bruelheide, H., Bonn, A., 2022. Temporal trends in the spatial bias of species occurrence records. Ecography 2022, 1–13 e06-219.
Boyd, R.J., August, T.A., Cooke, R., Logie, M., Mancini, F., Powney, G.D., Roy, D.B., Turvey, K., Isaac, N.J.B., 2023. An operational workflow for producing periodic estimates of species occupancy at national scales. Biol. Rev. 98, 1492–1508.
Boyd, R.J., Bowler, D.E., Isaac, N.J.B., Pescott, O.L., 2024. On the trade-off between accuracy and spatial resolution when estimating species occupancy from geographically biased samples. Ecol. Model. 493, 110739.
Boyd, R.J., Powney, G.D., Carvell, C., Pescott, O.L., 2021. occAssess: an R package for assessing potential biases in species occurrence data. Ecol. Evol. 11, 16177–16187.
Brooks, M., Rose, S., Altwegg, R., Lee, A.T., Nel, H., Ottosson, U., Retief, E., Reynolds, C., Ryan, P.G., Shema, S., Tende, T., Underhill, L.G., Thomson, R.L., 2022. The African bird atlas project: a description of the project and BirdMap data-collection protocol. Ostrich 93, 223–232.
Bull, P.C., Gaze, P.D., Robertson, C.J.R., 1985. The Atlas of Distribution of New Zealand Birds. Ornithological Society of New Zealand, Wellington.
Carvalheiro, L.G., Kunin, W.E., Keil, P., Aguirre-Gutiérrez, J., Ellis, W.N., Fox, R., Groom, Q., Hennekens, S., Van Landuyt, W., Maes, D., Van de Meutter, F., Michez, D., Rasmont, P., Ode, B., Potts, S.G., Reemer, M., Roberts, S.P.M., Schaminée, J., WallisDeVries, M.F., Biesmeijer, J.C., 2013. Species richness declines and biotic homogenisation have slowed down for NW-European pollinators and plants. Ecol. Lett. 16, 870–878.
Coelho, M.T.P., Barreto, E., Rangel, T.F., Diniz-Filho, J.A.F., Wüest, R.O., Bach, W., Skeels, A., McFadden, I.R., Roberts, D.W., Pellissier, L., Zimmermann, N.E., Graham, C.H., 2023. The geography of climate and the global patterns of species diversity. Nature 622, 537–544.
Craven, D., Sande, M.T., Meyer, C., Gerstner, K., Bennett, J.M., Giling, D.P., Hines, J., Phillips, H.R.P., May, F., Bannar-Martin, K.H., Chase, J.M., Keil, P., 2020. A cross-scale assessment of productivity–diversity relationships. Glob. Ecol. Biogeogr. 29, 1940–1955.
Dobson, A.D.M., Milner-Gulland, E.J., Aebischer, N.J., Beale, C.M., Brozovic, R., Coals, P., Critchlow, R., Dancer, A., Greve, M., Hinsley, A., Ibbett, H., Johnston, A., Kuiper, T., Le Comber, S., Mahood, S.P., Moore, J.F., Nilsen, E.B., Pocock, M.J.O., Quinn, A., Travers, H., Wilfred, P., Wright, J., Keane, A., 2020. Making messy data work for conservation. One Earth 2, 455–465.
Dorazio, R.M., 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Glob. Ecol. Biogeogr. 23, 1472–1484.
Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46.
Doser, J.W., Finley, A.O., Kéry, M., Zipkin, E.F., 2022. spOccupancy: an R package for single-species, multi-species, and integrated spatial occupancy models. Methods Ecol. Evol. 13, 1670–1678.
Doser, J.W., Finley, A.O., Kéry, M., Zipkin, E.F., 2024. spAbundance: an R package for single-species and multi-species spatially explicit abundance models. Methods Ecol. Evol. 15, 1024–1033.
Eichenberg, D., Bowler, D.E., Bonn, A., Bruelheide, H., Grescho, V., Harter, D., Jandt, U., May, R., Winter, M., Jansen, F., 2021. Widespread decline in central European plant diversity across six decades. Glob. Chang. Biol. 27, 1097–1110.
Essl, F., Moser, D., Dirnböck, T., Dullinger, S., Milasowszky, N., Winter, M., Rabitsch, W., 2013. Native, alien, endemic, threatened, and extinct species diversity in European countries. Biol. Conserv. 164, 90–97.
European bird census council (Ed.), 2020. European Breeding Bird Atlas. Lynx Edicions, Barcelona.
Farmer, R.G., Leonard, M.L., Mills Flemming, J.E., Anderson, S.C., 2014. Observer aging and long-term avian survey data quality. Ecol. Evol. 4, 2563–2576.
Faurby, S., Davis, M., Pedersen, R.Ø., Schowanek, S.D., Antonelli, A., Svenning, J., 2018. Phylacine 1.2: the phylogenetic atlas of mammal macroecology. Ecology 99, 2626.
Field, R., Hawkins, B.A., Cornell, H.V., Currie, D.J., Diniz-Filho, J.A.F., Guégan, J., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Turner, J.R. G., 2009. Spatial species-richness gradients across scales: a meta-analysis. J. Biogeogr. 36, 132–147.
Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., Zarzo-Arias, A., Barták, V., Moudrý, V., 2022. Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. Methods Ecol. Evol. 13, 2289–2302.
Gaston, K.J., 2003. The Structure and Dynamics of Geographic Ranges. Oxford Univ. Press.
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian data analysis, 3 edition. CRC Press.
Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4, 379–391.
Goury, R., Bowler, D.E., Harrower, C., Münkemüller, T., Vallet, J., Yearsley, J.M., Thuiller, W., Pescott, O.L., 2026. A practical guide to species trend detection with unstructured data using local frequency scaling (Frescalo). Ecography 2026, 1–14 e08270.
Grattarola, F., Bowler, D.E., Keil, P., 2023. Integrating presence-only and presence–absence data to model changes in species geographic ranges: an example in the Neotropics. J. Biogeogr. 50, 1561–1575.
Grattarola, F., Martínez-Lanfranco, J.A., Botto, G., Naya, D.E., Maneyro, R., Mai, P., Hernández, D., Laufer, G., Ziegler, L., González, E.M., da Rosa, I., Gobel, N., González, A., González, J., Rodales, A.L., Pincheira-Donoso, D., 2020. Multiple forms of hotspots of tetrapod biodiversity and the challenges of open-access data scarcity. Sci. Rep. 10, 22045.
Grattarola, F., Tschernosterová, K., Keil, P., 2024. A continental-wide decline of occupancy and diversity in five Neotropical carnivores. Glob. Ecol. Conserv. 55, e03226.
Guillera-Arroita, G., 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. Ecography 40, 281–295.
Hartley, S., Kunin, W.E., 2003. Scale dependency of rarity, extinction risk, and conservation priority. Conserv. Biol. 17, 1559–1570.
Hill, M.O., 2012. Local frequency as a key to interpreting species occurrence data when recording effort is not known. Methods Ecol. Evol. 3, 195–205.
Hsieh, T.C., Ma, K.H., Chao, A., 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol. Evol. 7, 1451–1456.
Hughes, A.C., Dorey, J.B., Bossert, S., Qiao, H., Orr, M.C., 2024. Big data, big problems? How to circumvent problems in biodiversity mapping and ensure meaningful results. Ecography 2024, 1–15 e07115.
Hurlbert, A.H., Jetz, W., 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proc. Natl. Acad. Sci. 104, 13384–13389.
Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., Jarvis, S., Lahoz-Monfort, J., Pagel, J., Pescott, O.L., Schmucki, R., Simmonds, E.G., O'Hara, R.B., 2019. Data integration for large-scale models of species distributions. Trends Ecol. Evol. 35, 56–67.
Isaac, N.J.B., Pocock, M.J.O., 2015. Bias and information in biological records: Bias and information in biological records. Biol. J. Linn. Soc. 115, 522–531.
Isaac, N.J.B., Van Strien, A.J., August, T.A., De Zeeuw, M.P., Roy, D.B., 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. Methods Ecol. Evol. 5, 1052–1060.
Jenkins, C.N., Pimm, S.L., Joppa, L.N., 2013. Global patterns of terrestrial vertebrate diversity and conservation. Proc. Natl. Acad. Sci. USA 110.
Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K., Mooers, A.O., 2012. The global diversity of birds in space and time. Nature 491, 444–448.
Johnston, A., Matechou, E., Dennis, E.B., 2023. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. Methods Ecol. Evol. 14, 103–116.
Jurasinski, G., Beierkuhnlein, C., 2006. Spatial patterns of biodiversity - assessing vegetation using hexagonal grids. Biol. Environ. Proc. R. Ir. Acad 106B, 401–411.
Keil, P., Belmaker, J., Wilson, A.M., Unitt, P., Jetz, W., 2013. Downscaling of species distribution models: a hierarchical approach. Methods Ecol. Evol. 4, 82–94.

Keil, P., Biesmeijer, J.C., Barendregt, A., Reemer, M., Kunin, W.E., 2011. Biodiversity change is scale-dependent: an example from Dutch and UK hoverflies (Diptera, Syrphidae). Ecography 34, 392–401.
Keil, P., Chase, J.M., 2019. Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. Nat. Ecol. Evol. 3, 390–399.
Keil, P., Jetz, W., 2013. Downscaling the environmental associations and spatial patterns of species richness. Ecol. Appl. 24, 823–831.
Keil, P., Pereira, H.M., Cabral, J.S., Chase, J.M., May, F., Martins, I.S., Winter, M., 2018. Spatial scaling of extinction rates: theory and data reveal nonlinearity and a major upscaling and downscaling challenge. Glob. Ecol. Biogeogr. 27, 2–13.
Keil, P., Schweiger, O., Kuehn, I., Kunin, W.E., Kuussaari, M., Settele, J., Henle, K., Brotons, L., Pe'er, G., Lengyel, S., Moustakas, A., Steinicke, H., Storch, D., 2012. Patterns of beta diversity in Europe: the role of climate, land cover and distance across scales. J. Biogeogr. 39, 1473–1486.
Keil, P., Wilson, A.M., Jetz, W., 2014. Uncertainty, priors, autocorrelation and disparate data in downscaling of species distributions. Divers. Distrib. 20, 797–812.
Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., Martí, D., Anton, M., Klvaňová, A., Kalyakin, M.V., Bauer, H.G., Foppen, R.P.B., 2020. European Breeding Bird Atlas 2: Distribution, Abundance and Change. European Bird Census Council & Lynx Edicions, Barcelona.
Kéry, M., Royle, J., 2020. Applied Hierarchical Modeling in Ecology. *Volume 2: dynamic and advanced models*, 1st edn. Elsevier, Cambridge.
Kéry, M., Royle, J.A., 2015. Applied Hierarchical Modeling in Ecology. Academic Press.
Kéry, M., Royle, J.A., 2016. Applied hierarchical modeling in ecology. Volume 1: Prelude and Static Models. Elsevier/AP, Academic Press Is an Imprint of Elsevier, Amsterdam Boston.
Klímová, B., Divíšek, J., Večeřa, M., Chytrý, M., Danihelka, J., Kaplan, Z., Novotný, P., Skokanová, H., Wild, J., Axmanová, I., 2025. Spatial patterns of species richness in the Czech flora: effects of sampling intensity, environment and landscape history. Preslia 97, 541–566.
Klinkovská, K., Glaser, M., Danihelka, J., Kaplan, Z., Knollová, I., Novotný, P., Pyšek, P., Řezníčková, M., Wild, J., Chytrý, M., 2024. Dynamics of the Czech flora over the last 60 years: winners, losers and causes of changes. Biol. Conserv. 292, 110502.
Kunin, W.E., 1997. Sample shape, spatial scale and species counts: implications for reserve design. Biol. Conserv. 82, 369–377.
Kunin, W.E., 1998. Extrapolating species abundance across spatial scales. Science 281, 1513–1515.
Lahti, T., Lampinen, R., 1999. From dot maps to bitmaps: atlas florae Europae goes digital. Acta Bot. Fenn. 162, 5–9.
Lele, S.R., Moreno, M., Bayne, E., 2012. Dealing with detection error in site occupancy surveys: what can we do with a single survey? J. Plant Ecol. 5, 22–31.
Lennon, J.J., Koleff, P., GreenwooD, J.J.D., Gaston, K.J., 2001. The geographical structure of British bird distributions: diversity, spatial turnover and scale. J. Anim. Ecol. 70, 966–979.
Leroy, F., Reif, J., Vermouzek, Z., Šťastný, K., Trávníčková, E., Bejček, V., Mikuláš, I., Keil, P., 2024. Decomposing biodiversity change to processes of extinction, colonization, and recurrence across scales. Ecography 2024, 1–14 e06995.
MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83, 2248–2255.
McInerny, G.J., Purves, D.W., 2011. Fine-scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. Methods Ecol. Evol. 2, 248–257.
Mellado-Mansilla, D., Midolo, G., Ortega-Solís, G., Reif, J., Grattarola, F., Craven, D., Leroy, F., Perrone, M., Šťastný, K., Bejček, V., Keil, P., 2026. Temporal span of biodiversity monitoring mediates the effects of area and environment. bioRxiv. https://doi.org/10.64898/2026.02.09.704769.
Mertes, K., Jetz, W., 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41, 1604–1615.
Meyer, C., Kreft, H., Guralnick, R., Jetz, W., 2015. Global priorities for an effective information basis of biodiversity distributions. Nat. Commun. 6, 8221.
Midolo, G., Clark, A.T., Chytrý, M., Essl, F., Dullinger, S., Jandt, U., Bruelheide, H., Argagnon, O., Biurrun, I., Chiarucci, A., Ćušterevska, R., De Frenne, P., De Sanctis, M., Dengler, J., Divíšek, J., Dziuba, T., Ejrnæs, R., Garbolino, E., Illa, E., Jentsch, A., Jiménez-Alfaro, B., Lenoir, J., Moeslund, J.E., Napoleone, F., Pielech, R., Rumpf, S.B., Sanz-Zubizarreta, I., Silva, V., Svenning, J., Swacha, G., Večeřa, M., Vynokurov, D., Keil, P., 2025. Six decades of losses and gains in alpha diversity of European plant communities. Ecol. Lett. 28, e70248.
Monsarrat, S., Boshoff, A.F., Kerley, G.I.H., 2019. Accessibility maps as a tool to predict sampling bias in historical biodiversity occurrence records. Ecography 42, 125–136.
Montràs-Janer, T., Suggitt, A.J., Fox, R., Jönsson, M., Martay, B., Roy, D.B., Walker, K.J., Auffret, A.G., 2024. Anthropogenic climate and land-use change drive short- and long-term biodiversity shifts across taxa. Nat. Ecol. Evol. 8, 739–751.
Moudrý, V., Šímová, P., 2012. Influence of positional accuracy, sample size and scale on modelling species distributions: a review. Int. J. Geogr. Inf. Sci. 26, 2083–2095.
Nekola, J.C., White, P.S., 1999. Special paper: the distance decay of similarity in biogeography and ecology. J. Biogeogr. 26, 867–878.
Orme, C.D.L., Davies, R.G., Olson, V.A., Thomas, G.H., Ding, T.-S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Owens, I.P.F., Blackburn, T.M., Gaston, K.J., 2006. Global patterns of geographic range size in birds. PLoS Biol. 4, e208.
Ortega-Solís, G.R., Šťastný, K., Bejček, V., Telenský, T., Mellado-Mansilla, D., Zárybnicky, J., Grattarola, F., Zárybnická, M., Vermouzek, Z., Voříšek, P., Leroy, F., Tietje, M., Soria, C.D., Padulosi, E., Trávníčková, E., Wölke, F.J., Keil, P., 2026. Opening a standardized, spatially contiguous biodiversity database collected over 40 years: Czech breeding bird atlases 1973—77, 1985—89, 2001—03, and 2014—17.
Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., Abrego, N., 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecol. Lett. 20, 561–576.
Pagel, J., Anderson, B.J., O'Hara, R.B., Cramer, W., Fox, R., Jeltsch, F., Roy, D.B., Thomas, C.D., Schurr, F.M., 2014. Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. Methods Ecol. Evol. 5, 751–760.
Peach, M.A., Cohen, J.B., Frair, J.L., 2017. Single-visit dynamic occupancy models: an approach to account for imperfect detection with atlas data. J. Appl. Ecol. 54, 2033–2042.
Pescott, O.L., 2026. Unifying occupancy-detection and local frequency scaling (Frescalo) models. Ecol. Model. 511, 111367.
Pichler, M., Hartig, F., 2021. A new joint species distribution model for faster and more accurate inference of species associations from big community data. Methods Ecol. Evol. 12, 2159–2173.
Pocock, M.J.O., Tweddle, J.C., Savage, J., Robinson, L.D., Roy, H.E., 2017. The diversity and evolution of ecological and environmental citizen science. PLoS One 12, e0172579.
Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A., McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). Methods Ecol. Evol. 5, 397–406.
Prajzlerová, D., Barták, V., Keil, P., Moudrý, V., Zikmundová, M., Balej, P., Leroy, F., Rocchini, D., Perrone, M., Malavasi, M., Šímová, P., 2024. The relationship between remotely-sensed spectral heterogeneity and bird diversity is modulated by landscape type. Int. J. Appl. Earth Obs. Geoinf. 128, 103763.
Prendergast, J.R., Wood, S.N., Lawton, J.H., Eversham, B.C., 1993. Correcting for variation in recording effort in analyses of diversity hotspots. Biodivers. Lett. 1, 39.
Preston, C.D., 2013. Following the BSBI'S lead: the influence of the atlas of the British flora, 1962–2012. New J. Bot. 3, 2–14.
Rahbek, C., 2005. The role of spatial scale and the perception of large-scale species-richness patterns. Ecol. Lett. 8, 224.
Rapacciuolo, G., Young, A., Johnson, R., 2021. Deriving indicators of biodiversity change from unstructured community-contributed data. Oikos 130, 1225–1239.
Robertson, M.P., Cumming, G.S., Erasmus, B.F.N., 2010. Getting the most out of atlas data. Divers. Distrib. 16, 363–375.
Rosindell, J., Hubbell, S.P., Etienne, R.S., 2011. The unified neutral theory of biodiversity and biogeography at age ten. Trends Ecol. Evol. 26, 340–348.
Sadoti, G., Zuckerberg, B., Jarzyna, M.A., Porter, W.F., 2013. Applying occupancy estimation and modelling to the analysis of atlas data. Divers. Distrib. 19, 804–814.
Schäfers, G., Ebersbach, H., Reimers, H., Körber, P., Janke, K., Borggräfe, K., Landwehr, F., 2016. Atlas der Säugetiere Hamburgs. *Artenbestand, Verbreitung, Rote Liste, Gefährdung und Schutz*. Behörde für Umwelt und Energie, Amt f. Naturschutz, Grünplanung und Energie, Abteilung Naturschutz, Hamburg.
Soria, C.D., Ortega-Solís, G.R., Wölke, F.J.R., Barták, V., Leroy, F., Tschernosterová, K., Bejček, V., Herrando, S., Mikuláš, I., Šťastný, K., Ueta, M., Voříšek, P., Keil, P., 2026. Spatial autocorrelation of species diversity and distributions in time and across spatial scales. Glob. Ecol. Biogeogr. 35, e70221.
Šťastný, K., Bejček, V., Mikuláš, I., Telenský, Tomáš, 2021. Atlas hnízdního rozšíření ptáků v České republice 2014-2017, 1. vydání. Aventinum, Praha.
Stojmenovic, M., Nayak, A., 2008. Measuring the related properties of linearity and elongation of point sets. In: Ruiz-Shulcloper, J., Kropatsch, W.G. (Eds.), *Progress in Pattern Recognition, Image Analysis and Applications* Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 102–111.
Storch, D., 2016. The theory of the nested species–area relationship: geometric foundations of biodiversity scaling. J. Veg. Sci. 27, 880–891.
Storch, D., Marquet, P.A., Brown, J.H., 2007. Scaling Biodiversity. Cambridge University Press.
Stroh, P.A., Walker, K.J., Humphrey, T.A., Pescott, O.L., Burkmar, R.J., 2023a. Plant atlas 2020: mapping changes in the distribution of the British and Irish Flora. Botanical Society of Britain and Ireland, Princeton University Press, Durham, Princeton, New Jersey.
Stroh, P.A., Walker, K.J., Humphrey, T.A., Pescott, O.L., Burkmar, R.J., 2023b. Plant Atlas 2020: Mapping Changes in the Distribution of the British and Irish Flora: volume 1–2. Princeton University Press, Princeton, New Jersey Oxford.
Svenning, J.-C., Skov, F., 2004. Limited filling of the potential range in European tree species: limited range filling in European trees. Ecol. Lett. 7, 565–573.
Szabo, J.K., Vesk, P.A., Baxter, P.W.J., Possingham, H.P., 2010. Regional avian species declines estimated from volunteer-collected long-term data using list length analysis. Ecol. Appl. 20, 2157–2169.
Telfer, M.G., Preston, C.D., Rothery, P., 2002. A general method for measuring relative change in range size from biological atlas data. Biol. Conserv. 107, 99–109.
Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehikoinen, A., De Jonge, M.M.J., Oksanen, J., Ovaskainen, O., 2020. Joint species distribution modelling with the R -package Hmsc. Methods Ecol. Evol. 11, 442–447.
Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234.
Ueta, M., Uemura, S., Hayama, S., Ara, T., Takagawa, S., Kawaji, N., Kajita, M., Nakamura, N., Tahata, S., Uchiyama, Y., 2023. The third dataset of "The breeding bird atlas of Japan".
Van Swaay, C.A.M., Warren, M.S., 1999. Red Data Book of European Butterflies. Council of Europe Publishing, Strasbourg.
Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B., Telfer, M. G., Jeffcoate, S., Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N.,

Moss, D., Thomas, C.D., 2001. Rapid responses of British butterflies to opposing forces of climate and habitat change. Nature 414, 65–69.
Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3, 385.
Witte, J.M., He, F., Groen, C.L.G., 2008. Grid origin affects scaling of species across spatial scales. Glob. Ecol. Biogeogr. 17, 448–456.
Xu, Q., Wang, X., Yi, J., Wang, Y., 2024. Bias correction in species distribution models based on geographic and environmental characteristics. Eco. Inform. 81, 102604.
Ziolkowski, D., Lutmerding, M., Aponte, V., Hudson, M.-A., 2022. 2022 release - north American breeding bird survey dataset (1966-2021).