Temporal changes in distributions and the species atlas: How can British and Irish plant data shoulder the inferential burden?

Oliver L. Pescott[1*], Thomas A. Humphrey[2], Peter A. Stroh[3] & Kevin J. Walker[4]

[1]UK Centre for Ecology & Hydrology, Wallingford, UK; [2]Botanical Society of Britain and Ireland (BSBI), Wallingford, UK; [3]BSBI, Cambridge, UK; [4]BSBI, Harrogate, UK

* **Corresponding author:** Oliver Pescott: oliver.pescott@ceh.ac.uk

Abstract
Species distribution atlases often rely on volunteer effort to achieve their desired coverage, an activity now typically discussed, at least in academia, under the general theme of "citizen science". Such data, however, are rarely without complex biases, particularly with respect to the estimation of trends in species' distributions over many decades. The data of the Botanical Society of Britain and Ireland (BSBI) are no exception to this, and both careful thought in data aggregation (spatial, temporal, and taxonomic) and appropriate modelling procedures are required to overcome these challenges. We discuss these issues, with a primary focus on the statistical models that have been put forward to adjust for such biases. Such models include the Telfer method, various "reporting rate" approaches based on generalised linear models, the frequency scaling using local occupancy ("Frescalo") model, occupancy models, and spatial smoothing methods. In each case the strengths and limitations in relation to estimating trends from distribution data with important time-varying biases are assessed. Various properties of BSBI data, in particular the increasing numbers of records at fine spatial and temporal scales over the past century, coupled with a general lack of re-visits to sites at such finer scales and the time-varying biases previously mentioned, imply that methods that can be sensibly applied at coarser levels are likely to be most appropriate for estimating accurate long-term trends in distributions. We conclude that Frescalo, which can be seen as a type of occupancy model where an adjustment for overlooked species is made in relation to spatial rather than temporal replication, whilst simultaneously adjusting for variable regional effort, is currently the most sophisticated tool for achieving this. Although recording community-accepted adjustments to data collection practices may allow for a greater application of occupancy modelling or other approaches in the future, methods that seek accurate trends over the long-term are necessarily limited either to scales at which various properties of the data in hand are most likely to be unbiased, or at which the biases are well enough understood to be modelled accurately.

Introduction

The collection of species' occurrence data is a fundamental activity within the science of ecology. Whether it is to increase our understanding of species' distributions at large scales, or to monitor abundances at much finer resolutions, the unit of knowledge representing the occurrence of a species at a particular point in space and time can be seen as one of the cornerstones of ecological understanding (Kéry & Royle, 2016). The desire to visualise and analyse such data at different scales of space and time has resulted in the publication of hundreds of species atlases across the globe, although this activity has been particularly associated with Europe and North America (Preston, 2013). Regional traditions of atlas production have also varied: British and Irish outputs have typically summarised information using a grid-based approach, whereas North American atlases have normally opted for administrative units, at least for plants (Preston, 2013). Preston (2013) provides a useful overview of various national and taxonomic trends in approaches to species atlases, with perhaps the main message being that such activities "defy generalisation", particularly in Europe.

In recent years the global expansion of interest in "citizen science", coupled with the ever-increasing power of information technology and the trend towards open data, have increased both the amount of data collected and the ability of analysts to access and explore such datasets (August *et al.*, 2015). The visualisation and analysis of species' distributions[1] across space and time remains a key goal of such activity (e.g. Dunn & Weston, 2008; Robertson, Cumming & Erasmus, 2010; Pescott *et al.*, 2015b), and various analytical approaches have arisen to cope with the particular forms in which such data come (Isaac *et al.*, 2014). Heterogeneity in the way in which atlas data have been collected by volunteers is a fact of life, and presents challenges for any analysis, with variation in the spatial, temporal, and taxonomic resolution of data all needing to be carefully considered (Pescott, Humphrey & Walker, 2018). Although there may be identifiable "best practices" for the analysis of such data in relation to particular general goals (Johnston *et al.*, 2019), inevitably any approach will also need to take into account features of the dataset in hand (Hill, 2012). New or very recently initiated projects may have the luxury of collecting data in a particular way in order to maximise certain types of information (e.g. see the recommendations of Altwegg & Nichols, 2019), but for analyses that wish to make use of less structured data covering many decades, then the decisions required for any analysis are likely to be manifold, and trade-offs may have to be explored (Robertson *et al.*, 2010; Pescott *et al.*, 2018). "Best practices" also imply top-down control of project design and data collection; this is not always possible if volunteer effort is partly coordinated by the volunteers themselves, as is the case for many British and Irish biological recording schemes and societies (Pocock *et al.*, 2015). As Allen (1976) pointed out many years ago, the success of

---

[1] The term distribution, or range, has been used to cover a variety of different measures (Maes *et al.*, 2015), including "area of occupancy" (AOO), "extent of occurrence" (EOO), and various site-level occupancy metrics (both indices thereof and true probabilities of occurrence). Note that within the current paper, a site is assumed to be a grid square at some scale, although this does not have to be the case.

natural history-based data collection in Britain and Ireland has, to a large extent, been the result of an "unconscious tradition of compromise" between professionals and amateurs. It should therefore not be forgotten that the unilateral imposition of particular approaches to fieldwork risks corroding the goodwill of the recording community, and reducing participation, if carried out with no regard to local values.

In this paper, we, (i), briefly review some features of British and Irish vascular plant data as collected in the field, digitised from herbarium specimens and Floras, or otherwise collated by members and associates of the Botanical Society of Britain and Ireland (BSBI),[2] that are particularly relevant to our aim of making reliable inferences about changes in species' distributions over time; and, (ii), our main focus, explore and critically assess the methods that are currently available for using this impressive data resource for analysing such changes at various scales. The overall aim is to outline the various strengths and weaknesses of the methods as they apply to these data, thereby focusing effort on the investigation of particular questions that may need answering in advance of future analyses, whether for Britain and Ireland or elsewhere.

British and Irish vascular plant data: A brief look at some patterns and biases

Pescott *et al.* (2018) have provided an overview of some of the key features of British and Irish plant occurrence data that must be taken into consideration for their analysis, and many of these points have also been made before by other workers (e.g. Rich, 1998; Telfer *et al.*, 2002; Rich & Karran, 2006; Hill, 2012; Groom, 2013). In presenting the figures and maps below, we merely wish to provide a small overview of some of the patterns and biases within occurrence data that the methods of the following sections seek to overcome in order to produce valid inferences regarding true distributional change; as such, these figures should not be taken as the last word on the BSBI distribution dataset for the forthcoming BSBI "Atlas 2020" (Walker *et al.*, 2010).

Recording activity waxes and wanes across time and space, as most human activities do, and, likewise, the results of such activities also differ depending on variables such as the skill and knowledge of the recorders, the time spent on a recording visit, and the detectability of the species sought. Such influences on datasets of biological records have been summarised by various researchers (e.g. Rich, 1998; Isaac *et al.*, 2014; Isaac & Pocock, 2015), and clearly illustrated in the introductory chapters of numerous atlases (e.g. Preston et al., 2002a; Preston, 2014; Preston & Rorke, 2014). A sometimes overlooked fact regarding some of this variation, at least among analysts, is that it is also influenced by changes in technology and cultural practices (Preston, 2014), as well as by "effort" per se. For example, Figure 1 below plots the numbers of records associated with different spatial grain sizes in the BSBI database over time within four different regions. It is extremely difficult to relate the number of records made in each category to changing recorder "effort" on the ground, because within different time periods particular philosophies and practices of atlas data collection were in play. The date-

---

[2] Although other datasets and collations exist for these areas (e.g. Carey *et al.*, 2008), the BSBI (http://www.bsbi.org) is the only organisation continually collecting vascular plant records across Britain and Ireland with the specific purpose of producing comprehensive distribution datasets.

classes covering the main recording periods for the two national vascular plant atlases (Perring & Walters, 1962 [1950–1969]; Preston *et al.*, 2002b [1987–1999]) show a preponderance of hectad (10 × 10 km square) records (Fig. 1). This is because during these projects, data-basing species occurrences for the purpose of producing maps largely involved master lists of species detected at this scale; however, variable numbers of actual field visits inevitably underlie the hectad summary lists for these periods. Likewise, a trend towards recording at the tetrad (2 × 2 km square) level appeared during the period 1987–1999, but is now being replaced by a preference for monad (1 × 1 km square) or finer-scale records (Fig. 1). As noted in several places below in our methodological discussions, this has implications for the choice of scale at which one might seek to apply any given method, and the potential biases that these decisions can introduce (Pescott *et al.*, 2018).

Data at finer spatio-temporal scales more clearly reflect actual field effort, and mapping the numbers of recording day visits, as inferred from occurrence record dates and resolutions, provides a somewhat clearer picture of activity in recent years (Figs 2c,d). However, data for earlier periods (Figs 2a,b) present a complex, compound, picture of such activity crossed with the way in which data were collected and/or databased within a region. Counties with few or no day visits at the tetrad scale or better in Figures 2a and 2b may reflect a low level or lack of recording activity, or the fact that data were only submitted for inclusion in the national database as summaries at some larger scale. The potential pitfalls in naively using such data for the creation of long-term distributional trends for species should be clear.

Whilst field effort and results are no doubt highly correlated, an alternative way of learning about patterns and potential biases in data is to examine the outcomes of recording, rather than the visit-level effort that went into achieving those outcomes (Fig. 3). Following Preston & Rorke (2014), Figure 3 compares two broad time periods in terms of the proportion of the overall species list for a hectad (1970–2019) that was recorded in each of two sub-periods (1970–1999 and 2000–2019). Such an approach highlights those areas that were well-recorded in a sub-period relative to the recording list for the overall period. Whilst such maps can be hard to interpret from an ecological point of view, big changes that are roughly equivalent to county areas (e.g. Cardiganshire in mid-Wales in Fig. 3) seem much more likely to represent changes in recorder activity than wholesale changes to a local flora (Preston & Rorke, 2014). The various spatio-temporal biases indicated by Figures 2 and 3, and the fact that these will also strongly vary with spatial scale over time given regional trends in the favoured spatial grain size for recording and databasing (Fig. 1), should be kept firmly in mind during the discussions that follow.

England
Hectad
Tetrad
Monad or finer

Ireland
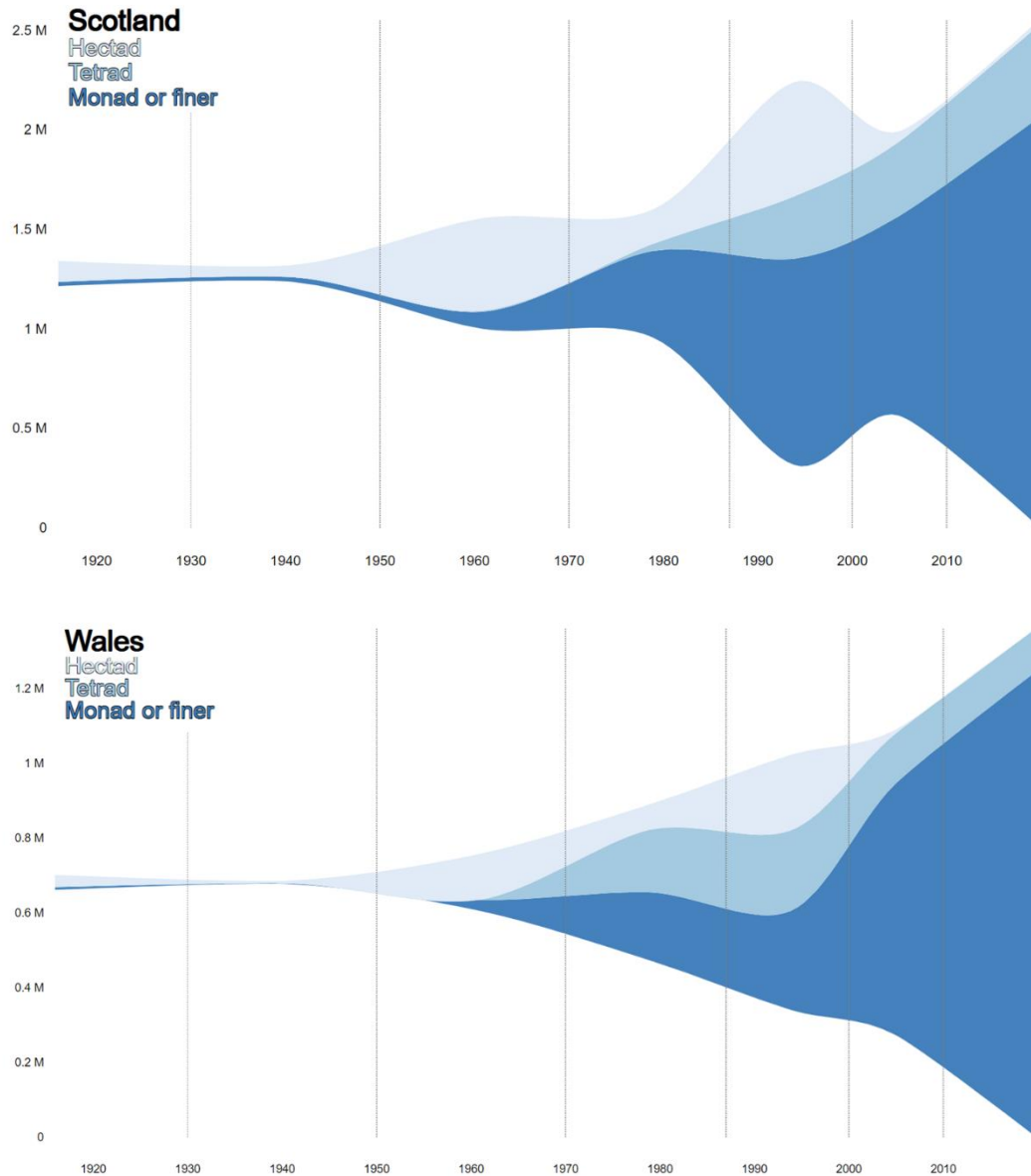Hectad
Tetrad
Monad or finer

254
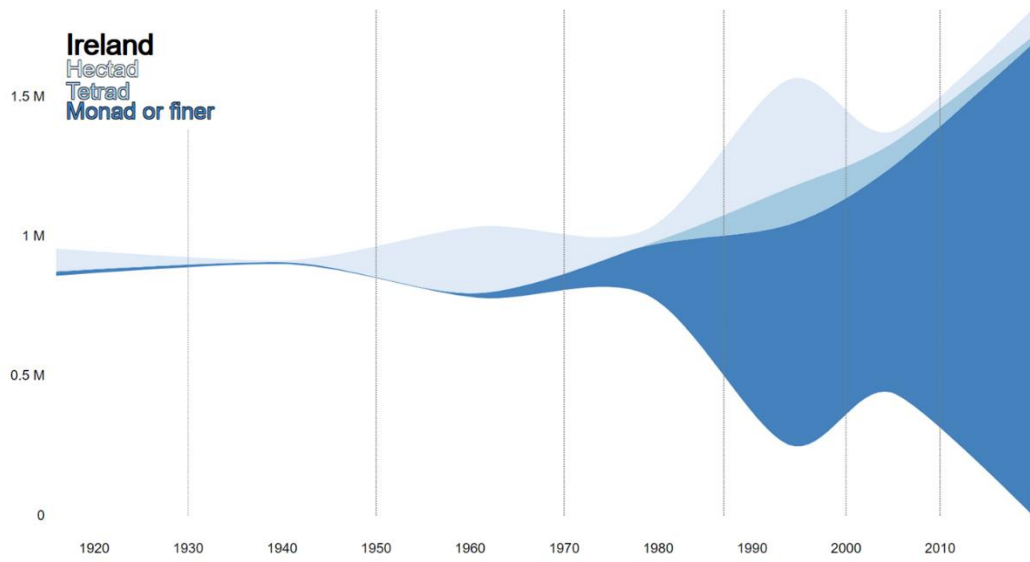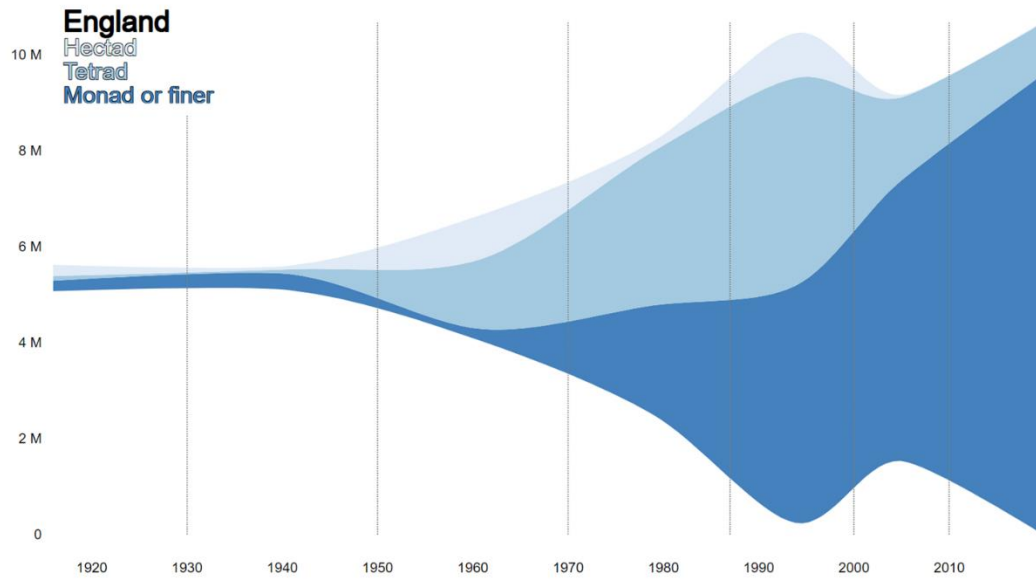
Figure 1. The numbers of records at different spatial scales per BSBI recording date-class, per region. Hectad = 10 × 10 km; tetrad = 2 × 2 km; monad or finer = 1 × 1 km or some finer grid-based resolution. Broken vertical lines represent BSBI date-class boundaries. The first date-class is all records pre-1930 (going back to the first records in the database), however, the overall counts for the period are here plotted arbitrarily for 1917 for ease of presentation. The other classes are: 1930–49; 1950–69; 1970–86; 1987–99; 2000–09; and 2010–19 (data from the BSBI Distribution Database (DDb) were accessed Oct. 2019). The actual numbers of records for each spatial scale are given at the mid-point of each class; the transitions between classes are smoothed interpolations only. The y-axis scales are different between regions; note also that in such "stream graphs" the y-axis intercept is arbitrary, as this axis is essentially only a scale bar for the stacked areas graphed. Plotted using the streamgraph package for R developed by Bob Rudis (2015; https://hrbrmstr.github.io/streamgraph/).
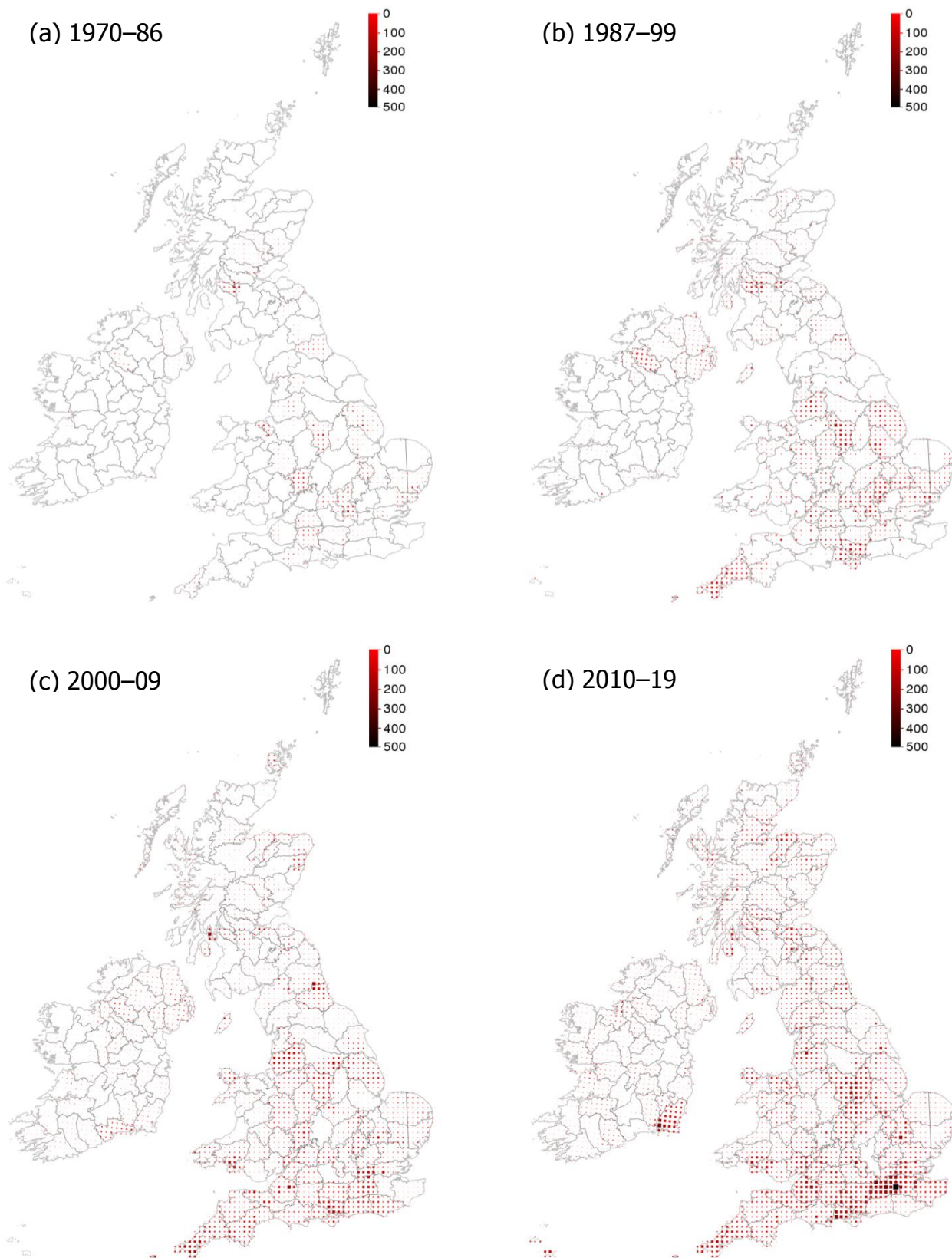
Figure 2. The number of recording days per hectad (10 × 10 km square), for each of the last four BSBI date-classes. A "recording day" is here defined as any single day where at least 40 records were made at the tetrad (2 × 2 km square) level or finer (data from the BSBI DDb accessed Oct. 2019). Hectad symbols are scaled, as well as coloured, according to the total number of recording days in the period. Colour and size scaling are uniform for comparability between maps. The background shows the British and Irish vice-counties (administrative recording units).

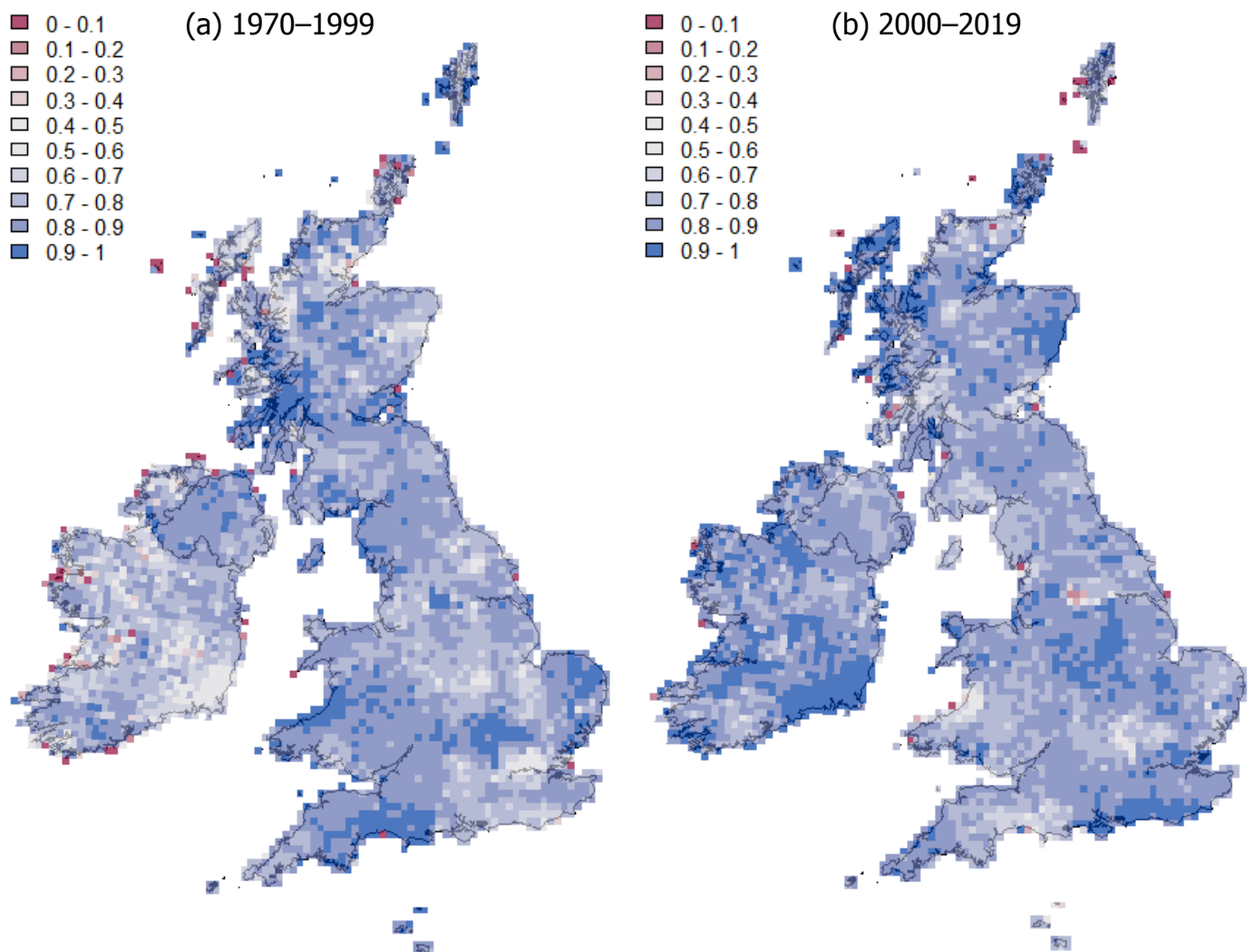**Figure 3. Number of vascular plant and charophyte species recorded from each hectad (10 × 10 km square) between (a) 1970–1999, and (b) 2000–2019, expressed as a proportion of the total number of such species recorded overall in the period 1970–2019 (data from the BSBI DDb accessed Oct. 2019).**

Analytical solutions?

Since the publication of the *New Atlas of the British & Irish Flora* (Preston *et al.*, 2002b), several new analytical methods for biological records made with variable recording effort (or outcomes) have been suggested (Isaac *et al.*, 2014). Foremost amongst these are perhaps the local frequency scaling, or "Frescalo", method of Hill (2012) and the occupancy modelling approach adapted for "opportunistic" recording by Kéry, van Strien and colleagues (van Strien *et al.*, 2010; van Strien *et al.*, 2013; Kéry *et al.*, 2010), although other options have also been put forward. We discuss these methods, as well as older techniques, in terms of their approaches, assumptions, strengths and limitations below.

The Telfer approach

Telfer's approach (Telfer *et al.*, 2002), developed for the *New Atlas* (Preston *et al.*, 2002b), models the relationship between counts of species' site occurrences across two different survey periods; the site counts in each period are expressed as logit-transformed proportions, with the same denominator (this being the total number of sites visited in both periods). Sites have typically been taken as 10 × 10 km squares (hectads), although the method could be applied at any spatial scale. The key assumption of this method is that changes in recorder effort affect all included species[3] equally across the area of interest. Once this assumption is made, the residuals from the regression of the logit-transformed proportions for period two on those for period one can be used as species-specific indices of change; "this index, though calculated as relative change in *recorded* range size, is proportional to relative change in *actual* range size" (Telfer *et al.*, 2002; original emphases).

The Telfer method attempts to overcome issues of variable geographic coverage and recording intensity by dealing only with sites visited in both periods, and by assuming that the relative recording frequency of the included species is unbiased; that is to say, that common species are still recorded more frequently than rare species, independently of overall changes in effort. This second assumption is likely to be reasonably robust across large areas, particularly if the time periods investigated have both been subject to atlas projects (as with the application of the method by Preston *et al.*, 2002b for example). However, the method is limited to separate comparisons of two periods; some very rare species may need to be completely excluded; and large scale changes that have actually affected a large proportion of the species included may be confounded with overall shifts in effort (Telfer *et al.*, 2002). In addition, some information may be wasted by only including sites that have been visited in both periods.

The Telfer method is one of several discussed in the current paper that were evaluated by Isaac *et al.* (2014) using a simulation approach. Isaac *et al.* (2014) made an impressive effort to explore the sensitivity of numerous methods for the

---

[3] Although we use the word species throughout for readability, this should be understood as referring to the designated group of species and species aggregates deemed appropriate for any particular analysis (Telfer, Preston & Rothery, 2002; Hill & Preston, 2015; Pescott *et al.*, 2018). For example, taxa with inconsistent taxonomic approaches across time periods are likely to be best treated at an aggregate level for many of the methods detailed here; likewise, critical taxa, hybrids, and many neophytes may often be best excluded due to uncertainties about the consistency and/or general representativeness of the recording effort giving rise to such records.

analysis of biological records against sets of simulated data, varied in different ways to represent types of recording bias likely to be present in such datasets. These simulations indicated that the Telfer method had a low type I (false positive) error rate, but also low power (i.e. a high rate of false negatives). That is to say, it appeared to be a conservative method when confronted with various types of simulated recording bias:[4] the method rarely concluded that a species had increased when it had not, but was also prone to not detecting true change, particularly in the presence of systematic trends in the detectability[5] of a species, or when similar trends applied across a large proportion of the species considered (Isaac *et al.*, 2014), both situations previously flagged as problematic by Telfer *et al.* (2002). However, it should also be pointed out that Telfer *et al.* (2002) did not advocate using their method with a view to determining whether any particular change was "significant", they instead noted that the continuous nature of their index circumvented issues with deciding on significance levels, a stance which now seems prescient in light of the increasing recognition of the noise and confusion that "significance" cut-offs can add to statistical reporting and inference (Amrhein *et al.*, 2019a; McShane *et al.*, 2019).

Reporting rate approaches
A much broader suite of approaches has developed around the idea of analysing the relative "reporting rate" of a species over time. The idea here is similar to the Telfer method, in that a relative index is used to account for variation in overall effort between time periods. The key assumption is again that the relative rate of reporting of a species is indicative of its true status, even if the overall effort has varied between time periods. Perhaps the simplest implementation of this approach is the application of a statistical test (e.g. the chi-squared test) to a 2 × 2 contingency table to determine whether the proportion of records of a particular species, out of all other records made, differs between two time periods (Desender & Turin, 1989). Note here that, due to the almost inevitable presence of duplicates within large databases of biological records, this type of test is typically best done on relative numbers of unique site × date × species occurrences (normally referred to as "visits" if the dates are resolved to a single day), rather than on raw record counts. The reporting rate approach has been used in a wide range of studies, continuing up to the present time (e.g. Desender & Turin, 1989; Seaward, 1998; Ball *et al.*, 2011; Roy *et al.*, 2012; Hill & Preston, 2014; Pescott *et al.*, 2015a; Preston & Hill, 2019; Boersch-Supan *et al.*, 2019). The approach can be generalised to a binomial generalised linear (mixed) model, allowing one to adjust for other covariates, such as indices of recording effort (e.g. Ball *et al.*, 2011), and/or take random effects, such as sites or regions, into account as deemed appropriate (e.g. Roy *et al.*, 2012). This linear modelling approach can be used on aggregated data (number of detections per number of visits) from broad time periods (Ball *et al.*, 2011), or applied to visit-level disaggregated data on individual species detections and non-detections; this latter approach opens up the option of using visit-level covariates to

---

[4] See Table 1 of Isaac *et al.* (2014; an open access paper) for an overview of the recording scenarios used.
[5] By detectability we mean the chance that a species is recorded if present. This is not a constant feature of a species, but will also depend on the time spent searching, the extent of the area searched, recorder expertise, the time of year etc.

adjust for some elements of recording effort, for example, the number of species recorded during a visit, known as the "list length" (Franklin, 1999; Szabo *et al.*, 2010). Finally, as with the Telfer method discussed above, so can reporting rate approaches apply criteria for the inclusion of particular sites, visits or time periods to any analysis (Ball, 2010; Roy *et al.*, 2012; Boersch-Supan *et al.*, 2019).

One key extension to the reporting rate approach as applied to broad time periods was that of Latour & van Swaay (1992); these authors noted that this approach would be more robust if the relative rate was reported only in relation to the average site count of those species that were considered to have been common and stable throughout the time periods analysed. That is, the "average" common species is used as an index of effort within periods. Maes & van Swaay (1997) provided the details of the criteria suggested for pre-selecting these so-called "benchmark" species against which to produce relative reporting rate trends for other species. In their formulation, species that were at least moderately common (occurring in ≥ 10% of their sites), and which showed a relatively flat time trend (-1 < regression slope b < 1), with a moderate to high R^2 of ≥ 0.2 (indicating low variance around the trend), in their reporting rate were selected as benchmarks.[6] The use of benchmark species to estimate recording effort was developed by Hill (2012), whose separate method is discussed below.

As for the method of Telfer *et al.* (2002), reporting rate methods are best applied to datasets where geographic coverage is relatively similar across the periods analysed; uneven recording across space and time could lead to misleading results, as the reporting rate could then merely reflect different species' frequencies in different areas independently of actual trends over time. Note that the use of random effects in mixed models, for example specifying a random intercept for each site or region (Roy *et al.*, 2012), will not necessarily ameliorate this. The simple fact is that any reliable inference about change will be dependent on the assumption that each set of sites is representative of the wider population of interest; if they are unrepresentative, then the resulting trend estimate is likely to be badly biased. It might perhaps be unlikely that the types of sites included in analyses will change completely between time periods of interest,[7] but it is worth being clear about the issues that particular enhancements to models can and cannot fix: including a random intercept for sites, for example, allows similarities between data points collected within and between sites to be modelled (thus overcoming some issues relating to variable sample sizes; Gelman & Hill, 2007), but it does not overcome fundamental issues such as the potential confounding of site or visit types with time.[8]

---

[6] Note that such an approach may be less relevant when information-rich visit-level information is available; in this case the availability of data such as whether the recorded list was thought to be complete, the length of time spent recording etc., may encode more information on effort than such benchmarking (e.g. Boersch-Supan *et al.*, 2019). However, this will of course be dataset specific.

[7] Although complex interactions between spatial scale, time period, and a species' rarity do exist to trip up the unwary. For example, the vascular plant species recorded at small scales (< 1km^2) in the BSBI Database are increasingly likely to be an unrepresentative subset of rare, scarce, critical, or otherwise unusual taxa as one goes further back in time (Pescott *et al.* 2018).

[8] For a recent and comprehensively examined example of such a phenomenon, see Preston & Hill's (2019) bryophyte Flora of Cambridgeshire, where substantial changes in the types of sites visited, particularly affecting species data from churchyards and arable fields, were seen over time.

Isaac *et al.* (2014) tested a number of variations on the basic reporting rate approach, including options that used site filtering before model fitting, the use of list length as a covariate at the visit level to adjust for recording effort, and the use of a site random effect. A model including all of these features was also included in their simulation tests. Their investigations provided a number of useful insights. For example, their "control" scenario, where all simulated species had stable distributions, indicated that reporting rate approaches without a site effect had inflated type I error rates relative to the other methods investigated ("significant" trends were found at around twice the rate at which they were expected under the correct null model). This was also the case when the number of visits increased over time. In this result we can appreciate the impact of including a site effect to account for the fact that data points within a site are not independent. Reporting rate methods also performed poorly when a focal species of interest became more "detectable" over time; this is not unexpected, given that none of the suggested modifications to the basic reporting rate method directly account for this type of bias. Including a visit-level list length covariate controlled the type I error rate relative to other reporting rate models when a trend in list length was the main type of bias simulated, however, these models had low power to detect true trends (again, a conclusion that makes intuitive sense, given that reducing the number of species recorded on visits over time degrades the information content of the reporting rate statistic that is the key parameter enabling inference concerning such trends). More detail could be extracted from the results of Isaac *et al.* (2014), but the key message is that reporting rate methods, even with a variety of enhancements, are unlikely to provide full protection against the various biases that we may expect to find in datasets of biological records.[9] A more fundamental criticism of such approaches is that they actually model the product of occupancy and detectability (Kéry, 2011), rather than separating out the underlying ecological truth from the observation process that generated the data at hand (see the following two sections for more on this point).

Frequency scaling using local occupancy (Frescalo)
Mark Hill's (2012) "frequency scaling using local occupancy" method, also known as "Frescalo", takes the insight of Latour & van Swaay (1992), that the use of benchmark species is potentially a powerful way of estimating recorder effort, and develops its application across multiple local areas within a larger territory. One of the reasons for dividing a larger territory into smaller areas for the estimation of

---

[9] It is worth noting that, in those circumstances where the "full" reporting rate method, i.e. a method with site filtering, a list length covariate, and a random site effect, fails (for example, in the Isaac *et al.* (2014) "MoreDetectable" scenario, where the species of interest becomes more detectable over time), then the addition of covariates to the model could potentially solve this for any given species (e.g. an indicator covariate specifying whether a particular identification guide was available at a given point in time); however, this type of information could be confounded with time-based random effects, and would be difficult to quantify at the visit level (does the ID tool need to be present, or has the new knowledge passed into the recorders' minds?) Overcoming the "NonFocalDeclines" bias scenario of Isaac *et al.* (2014) using the reporting rate method is more challenging, as any method based on relative frequency will struggle to distinguish true increases in a species of interest from declines in sets of other species without independent measures of absolute change for calibration purposes (Telfer *et al.* 2002); on the other hand, in this scenario a random site effect does appear to provide reasonable control of error rates (type I and II), presumably because declines simulated for particular species are ultimately site-based, and a site-specific intercept models these correlations.

effort is that this may vary through space, as may appropriate benchmark species. Being able to adjust for variation in effort across space also goes some way towards overcoming the issue discussed above, where analyses of change may be undermined by mismatches in the overall recording footprint between time periods; although, of course, some data are still required to make these adjustments across a larger area—there is no absolutely free lunch. These local areas are termed "neighbourhoods" by Hill (2012), and in his original formulation were defined as a 100 hectad subset of the 200 hectads surrounding any particular focal hectad-site. The 100 hectad subset is selected based on a combination of each hectad's geographic distance from, and biological similarity to, the focal hectad. Therefore those hectads closest and most environmentally similar to the focal site will receive the largest weights (Fig. 4).

These neighbourhood weights are subsequently used in the calculation of weighted frequencies for all of the species in the neighbourhood; those species which are both frequent and in the most similar neighbourhood hectads therefore have the largest local weighted frequencies (a species that was present in all hectads in a neighbourhood would receive the maximum possible weighting). The mean of all these species weighted frequencies in any given neighbourhood is therefore a type of weighted mean (it is labelled the "frequency-weighted local mean [species] frequency" in Hill, 2012).

Perhaps the key insight behind the method is that this weighted mean frequency, $\sum_j f_{ij}^2 / \sum_j f_{ij}$, calculated across the j species in neighbourhood i, can be rearranged to give

$\sum_j f_{ij} / \left(\left(\sum_j f_{ij} \right)^2 / \sum_j f_{ij}^2 \right)$, an illuminating result for the quantitative ecologist, given

that the denominator $\frac{\left(\sum_j f_{ij} \right)^2}{\sum_j f_{ij}^2}$ is the reciprocal of the well-known Simpson's diversity

index, and one of the "effective species numbers" defined by Hill (1973). Thus, the weighted mean frequency in a neighbourhood can be understood as the average species richness of the hectads in the neighbourhood, $\sum_j f_{ij}$, divided by Hill number 2 (often written as N_2). Hill numbers are often called "true diversities" because they obey a set of principles that directly relate to our intuitive understanding of the diversity of a system, unlike entropy measures (Jost, 2006). The important point for the argument underpinning the Frescalo method is that true diversities like N_2 are multiplication invariant; that is to say, changes to a community that leave the relative frequencies of the species therein unaltered do not change the value of N_2. This means that $\frac{\sum_j f_{ij}}{N_2}$ can be seen as a measure of sampling intensity, because if a proportion p of records were deleted from a neighbourhood such that each species now had local frequency $p \times f_{ij}$, the average species richness of hectads in the neighbourhood $\sum_j f_{ij}$ would change by a factor of p, but N_2 would remain at the same value. In this sense, Hill (2012) suggests that $\sum_j f_{ij} / N_2$ can be used as an index of the sampling effort in a neighbourhood. This theoretical argument is supported by the empirical demonstration that weighted frequency curves for bryophyte species across five randomly selected neighbourhoods in Britain and Ireland could be shown to exhibit almost identical forms after adjustment for both local species richness variation and recording effort (Hill, 2012). This suggests a potentially quite general pattern at this scale, and a method that can be applied to

any taxon group across any territory where there has been at least enough survey effort to provide reasonable estimates of the relative frequencies of the species of interest.

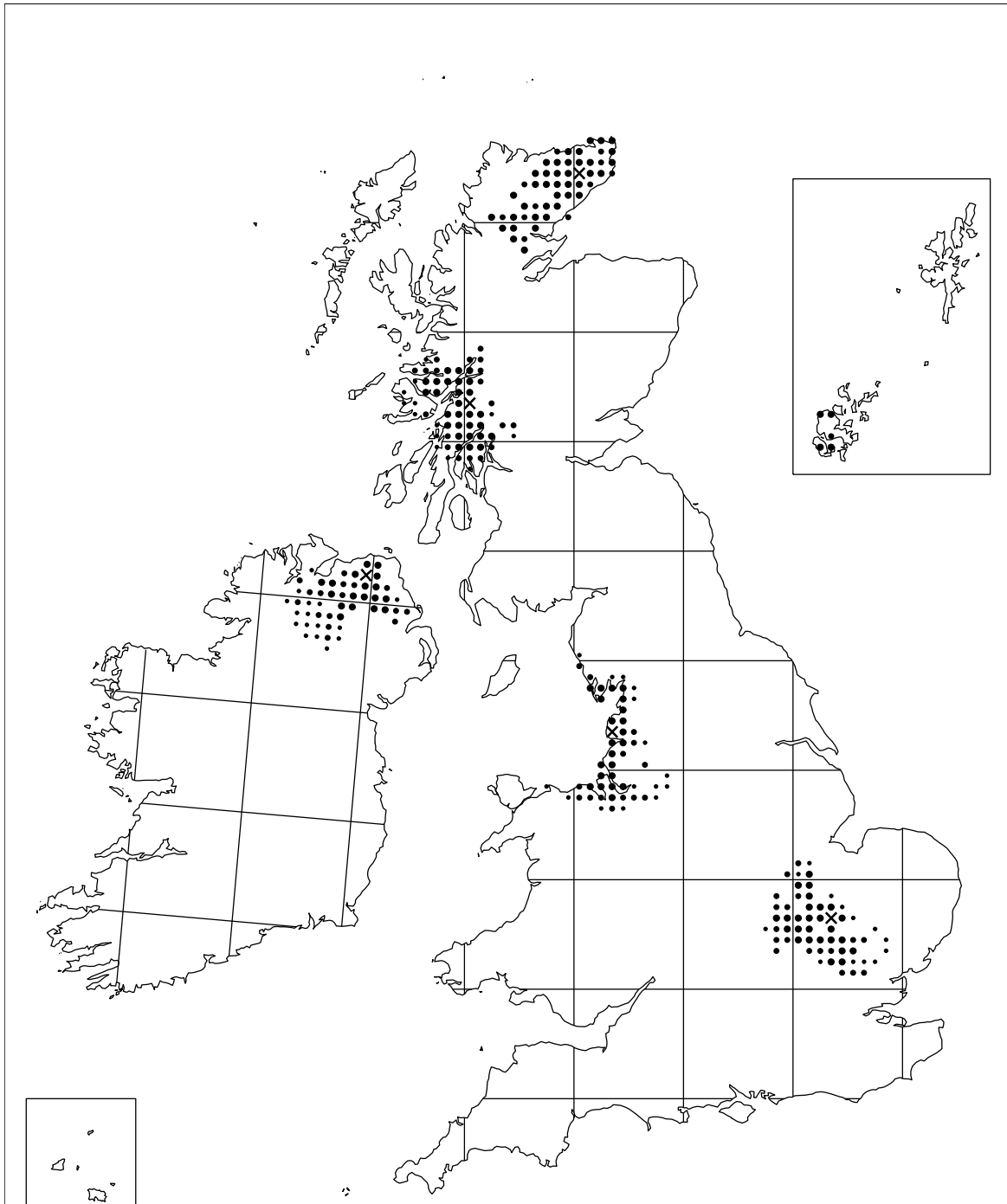**Figure 4. The weights for five randomly selected neighbourhoods in Britain and Ireland. In each case the focal site or hectad is indicated by a ×, other sites in each neighbourhood are indicated by a dot proportional to its weight, as calculated based on its geographical and environmental distance from the focal hectad. Figure originally prepared for Hill (2012) by Dr Colin Harrower (UKCEH).**

This pattern being the case, if a universal frequency-weighted local mean frequency value is specified in advance (labelled Φ (phi) by Hill, 2012), values of a sampling effort multiplier that adjust the observed mean weighted frequency in each neighbourhood to this standard value can be estimated. This sampling effort multiplier is defined as the rate parameter of a standard Poisson process modelling the probability of species discovery within each neighbourhood. Hill (2012) also uses the Poisson process to investigate change in a species local frequency between time periods. The probability[10] that a species j is recorded in hectad i in time period t is modelled as $P_{ijt} = 1 - \exp(-Q_{ijt}x_{jt})$, where x_{jt} is the time-dependent probability of finding species j, and Q_{ijt} is an estimate of recording effort and any time-independent element of the discovery of species j in hectad i. Estimating the values of Q_{ijt} so that x_{jt}, the time factors of interest, can be calculated then becomes the challenge. Hill (2012) uses the suggestion of Latour & van Swaay (1992) to use benchmark species as an index of recording effort to overcome this problem. After the standardisations for effort and richness noted above, Hill uses the top 27% of the expected number of species in the neighbourhood, across all time periods, as his benchmark set (see the original paper for more detail on this choice and related sensitivity analyses); also recall that these are local benchmarks chosen within each neighbourhood, overcoming the fact that the common species of one region may be quite different to those of another.[11] This approach has been used to analyse species' occurrence data in a number of studies (Balmer *et al.*, 2013; Fox *et al.*, 2014; Hill & Preston, 2014, 2015; Woodcock *et al.*, 2014; Pescott *et al.*, 2015a; Dyer *et al.*, 2017; Preston & Hill, 2019; White *et al.*, 2019), and has also been shown to perform well in tests using simulated data exhibiting various types of bias (Isaac *et al.*, 2014).

A key assumption of Frescalo is that local neighbourhoods should be surveyed in such a way that species' relative frequencies are reasonably well estimated, e.g. if only half the hectads in a neighbourhood have been surveyed, then common species should still be relatively more common than occasional species etc. in the dataset. Clearly this is an approximation, and its truth or otherwise will depend on the extent of local spatial and taxonomic biases in the dataset at hand. However, it should hold reasonably well for time periods that have seen semi-structured atlasing activity, particularly if we consider that the analysis is repeated across every neighbourhood in the wider territory considered, and therefore that a small proportion of slightly biased neighbourhoods may not have a large effect on the average result for a

---

[10] Note that the general expression $1 - e^{-\lambda x}$ is the cumulative distribution function of the exponential distribution; this defines the waiting time until an 'event' under a standard Poisson process. This relationship also underpins a link between the intensity of a Poisson distribution (λ) and occupancy probability (Kéry & Royle, 2016 pp. 110–112).

[11] Groom (2013) includes a reference to Hill (2012) within his general criticism that many methods for the analysis of unstructured data must "assum[e] that there has been no change in the occupancy of common species". Whilst it is the case for Frescalo that the neighbourhood sets of benchmark species are fixed across the time periods analysed, Hill (2012) noted that in his sensitivity tests using bryophyte distribution data, the exclusion of strongly increasing or decreasing species made very little difference to the estimated species' trends. Whilst this is not definite proof that this assumption is always inconsequential, coupled with the fact that the proportion of species used as benchmarks made no difference to estimated relative trends (Hill, 2012), the implication is that the recording effort information contained in such sets of species across space and time is fairly robust (at least when applied at scales where wholesale changes in land covers are unlikely to have taken place).

species in a particular time period. The assumption of unbiased relative recording frequencies in relation to actual frequencies is also an assumption of most of the other methods discussed here, and is not unique to Frescalo, although the fact that it is required to hold across multiple smaller areas for Frescalo does distinguish it from the Telfer method, and from applications of the reporting rate method at larger scales or across larger areas. Careful applications of Frescalo have considered the potential for particular data subsets to introduce large biases into analyses due to the fact that a particular tranche of recording activity was purposefully biased towards a particular habitat or set of species: for example, Hill & Preston (2014) investigated the effects of excluding one large Scottish woodland dataset in their analysis of national bryophyte trends, whilst Preston & Hill (2019) excluded the records of one particular recorder whose focus on liverworts, but not mosses, introduced a bias that was otherwise hard to account for when analysing changes in the bryophyte flora of Cambridgeshire. These types of decisions, however, could be made in relation to any analytical approach, and are a function of familiarity with the data being modelled, rather than being a special feature of the application of Frescalo. Finally, and as with other methods, decisions will have to be made regarding the taxonomic entities that are most appropriately modelled using this approach (see footnote 3).

Isaac *et al.* (2014) included two applications of Frescalo in their simulation study: first the approach was applied to annual time slices of their simulated dataset, and second after amalgamating these data into two time periods. The two period application of Frescalo gave similar results to the use of the method on an annual basis, albeit with a lower type I error rate and lower power. The annual application gave close to nominal type I error rates across most of the bias scenarios of Isaac *et al.* (2014), although biases that ultimately meant that the relative frequencies of species were not well represented (e.g. visiting sites because they had particular species, or species becoming relatively more detectable over time), resulted in higher false positive rates, as would be expected. Across most of the bias scenarios, the use of Frescalo annually had lower power than most of the other approaches tested, including the reporting rate approaches and occupancy modelling. Isaac *et al.* point out that this is because their simulation study is designed to mimic visit-based data at small spatial scales, whereas Frescalo was designed to be implemented at larger spatial and temporal resolutions. This means that the within-year visit-level information in the simulated datasets of Isaac *et al.* was unused by Frescalo. As an example, in the "MoreDetectable" bias scenario of Isaac *et al.*, the per visit detection probability of the *decreasing* focal species *increases* over time; at the level of the site × year summary these factors are likely to cancel each other out, therefore it is to be expected that Frescalo, with access only to the final species lists per site × year combination, almost always fails to detect the true underlying decline of the focal species in this scenario. Whether or not this is perceived as a weakness of the method can only be judged in relation to the dataset at hand: clearly, using Frescalo on a dataset consisting of highly informative visit-level data would be wasteful, but for a heterogeneous dataset containing various biases relating to the collection of visit-level data (e.g. visit-level information at fine spatial scales only exists for rarer species in early time periods), then the use of Frescalo after appropriate spatio-temporal aggregation may well be the more sensible strategy for avoiding biased inference.

Finally, Isaac *et al.* (2014) express the reservation that Frescalo requires the "user to make a variety of choices", and that these have a "considerable impact on the trend estimates that are produced". Whilst this may be the case to some degree, one of the choices specifically highlighted by Isaac *et al.* (2014), the selection of the proportion of species used as benchmark species locally, was explored using sensitivity analyses in the appendix of Hill (2012) and was there shown to have little to no impact on the estimation of relative trend patterns for a species (also see footnote 11). The other reservation of Isaac *et al.* (2014) concerning the definition of neighbourhoods used (Fig. 4), could rather be seen as a positive feature of the method, given that it underpins the ecological appropriateness of the species frequency curve approach taken. Careful choices imply careful thought, particularly in the face of biased data, and the requirement for preliminary thought in advance of applying any particular method is something to be strongly encouraged (Pescott *et al.*, 2018). Finally, we feel that it is also useful to consider that there are some commonalities between Frescalo and the occupancy modelling approach, and these are discussed below in the 'Challenges and limitations' part of the following occupancy modelling section.

Occupancy modelling
Occupancy, or site-occupancy, modelling was developed to overcome the fact that the vast majority of presence/absence data collected by naturalists and ecologists are in truth detection/non-detection data. False-negative errors, that is, overlooking a species during a survey that was in fact present, are clearly widespread. Unlike the reporting rate approaches described above, which actually model the product of true occupancy and detectability (i.e. detection/non-detection), confounding these parameters, occupancy models use a hierarchical approach in an attempt to clearly separate these two aspects of the recording process (MacKenzie *et al.*, 2006; Kéry & Royle, 2016). This approach can be thought of as using two linked Bernoulli generalised linear models, one that models true occupancy, and a detection model that models the conditional probability that a species that is present is actually detected (Kéry & Royle, 2016). Both models can incorporate fixed and random effects as per any standard regression modelling approach. In its most frequently encountered form, the logic of the approach rests on the availability of repeated visits[12] to a site to estimate a species' detectability within a time window (normally called a "closure period" in the occupancy modelling jargon). For example, if a species is present on a site, but only detected on a small fraction of the total number of visits within a closure period, then that information is used to adjust the overall estimate of a species' frequency: the model adjusts for the fact that a species exhibiting a low detectability is likely to be frequently overlooked. This method is based on the assumption that, within the closure (i.e. time) periods and at the spatial scale analysed, a species' detectability can be estimated from recorder visit data (Kéry *et al.*, 2010). The following three assumptions fundamental to the standard application of the method are set out by Kéry & Royle (2016). (1) The

---

[12] In analyses of unstructured (opportunistic) data, records made at a site (however defined) on a particular day are normally taken to indicate a single visit. However, for designed experiments (or through coincidence), different individuals or groups recording the site on the same day could also be considered as separate visits if they provide information regarding species' detectabilities.

"closure" assumption. The true occupancy of a species should be stable within the temporal and spatial grain at which the model is applied. We cannot estimate the actual occupancy of a species if, in fact, its true state at a site varied across the replicate visits from which we try to estimate the truth. (2) No false positives. The standard model assumes that there are no mistaken identifications of a species when it is truly absent. (3) Homogeneity of the detection probability. The detection probability of a species may vary between sites, and the inclusion of covariates in the detection model may not always explain this well. Although work on this topic suggests that this can often be modelled (Royle, 2006), the same study also warned that there are circumstances in which multiple models may fit the data equally well, but which provide different estimates of occupancy. Other assumptions that apply more generally to statistical models also apply (e.g. independence of samples, assumptions about the suitability of the distributions used to model variables), as for all the methods discussed here, but we do not discuss these further (see Kéry and Royle 2016 for more information).

The fact that both the "true" occupancy and detection parts of an occupancy model can include random effects and covariates makes many elements of model specification similar to the reporting rate approaches discussed above. For example, van Strien *et al.* (2013) included a list length variable in their detection model to better account for variable effort between surveyor visits, and Isaac *et al.* (2014) used an occupancy model with a list length covariate on the detection model, and added a random site effect to the true occupancy part of the model in their simulation tests. As with the reporting rate models, any number of important covariates could potentially be included in either part of the model; for example, the time of year, species' traits, and whether a species is flowering are all likely to be related to detectability to some extent (Kirby *et al.*, 1986).

A considerable literature on occupancy modelling exists, with a large part of this focusing on exploring the impacts of unmet model assumptions, and on enhancements of the basic model to account for these (see Kéry & Royle, 2016). Here we focus only on those issues that we consider particularly apply to the use of occupancy models for British and Irish "opportunistic", or semi-structured (Pescott *et al.*, 2015b), plant atlas data. We divide these into two general topics: (1) biases in the detection model; and (2) lack of repeat visits within the specified closure periods.

1. **Biases in the detection model.** If estimates of a species' detectability are biased in some way, then occupancy probabilities will be biased. Kéry & Royle (2016, p. 559) provide the following example: "if multiple surveys are only undertaken at the "better" sites, where [a species] density and therefore detection probability (p) may be higher on average, the resulting estimate of p will be biased high with respect to all sites and therefore the occupancy estimator will be biased low". A similar logic pertains to the example where rare species might disproportionately feature on short lists within a dataset, and some form of recorded list size is used as a covariate that is thought to be related to detectability as an index of search effort. If there is actually a negative relationship between list length and the probability of a species being observed (because it has been the target of special, focused, surveys), then detectability may be higher for short lists, again biasing occupancy

estimates downwards. This is despite the fact that, all else being equal, a 'random search' model of recording at the 1 km^2 scale implies that rare plant species should have low detectability, and so typically only feature on longer lists that are taken to indicate a greater search effort. Such biases could also vary with time and/or space e.g. an alien species might be more frequently recorded on shorter lists when it is a new arrival to an area.

2. **Lack of repeat visits within closure periods.** If, across a dataset for a species, only small numbers of repeat visits exist within closure periods, then occupancy models may have problems separating occupancy and detectability (Welsh *et al.*, 2013). This may not result in biased estimates of occupancy, in terms of long-run estimates for a given set of parameters, but it will increase the variance in these estimates (Welsh *et al.*, 2013; Guillera-Arroita *et al.*, 2014). This topic has been the subject of some debate in the literature and elsewhere (Welsh *et al.*, 2013; Welsh, Lindenmayer & Donnelly, 2015; Guillera-Arroita *et al.*, 2014; McGill, 2014). The debate ultimately comes down to whether one is happy to model detection/non-detection in a "naïve" way (e.g. using a reporting rate model), and accept an occupancy estimator that is highly likely to be biased (if the detection probability is less than perfect), but which is expected to have lower overall error[13] due a lower variance; or, to use an occupancy model with potentially inadequate data in order to recognise that little can be said about "true" occupancy, therefore representing the uncertainty underlying such estimates transparently (assuming all the other assumptions of a model are met). A comprehensive simulation study presented by Guillera-Arroita *et al.* (2014) showed that, with sites receiving only two visits within closure periods, naïve models that do not model occupancy and detection separately often have a lower mean-squared error than occupancy model estimates. Given that researchers normally only have one dataset, this means that in these situations an occupancy estimate from a naïve model is more likely to be closer to the true value, even though model estimates will be biased in a "long run" or average sense. It is not difficult to appreciate that this modelling decision will often be context-specific: one cannot request that atlas data from the distant past are re-collected!

The closely related issue of not having a high proportion of sites with more than one visit within a closure (time) period has not been rigorously explored to our knowledge: most research on occupancy model performance and design has tended to investigate the number of repeat visits as a constant across sites and years (e.g. the impact of all sites having two visits in each closure period versus all sites having five visits; Kéry & Royle, 2016). However, the fact that the performance of occupancy models is known to decline as the amount of information about detectability in a dataset declines (van Strien *et al.*, 2010) strongly implies that fitting occupancy models to data with a low overall proportion of sites with repeat visits within closure periods may often be a risky business. Although Kéry *et al.*

---

[13] Often measured as mean-squared error (MSE).

(2010) found that occupancy models apparently provided good trend estimates using opportunistic data on four bird species from the Swiss bird observation database, in that situation the "worst case" species in terms of data availability, for the Rock Thrush *Monticola saxatilis*, still had 42% of its 1 km site × year combinations with more than a single visit (and often the number of visits within such site × year combinations was considerably larger than two). Likewise, van Strien *et al.* (2013) demonstrated frequent congruence between dragonfly and butterfly occupancy trends derived from opportunistic data,[14] whilst accounting for variation in detectability due to the proportion of species reported by recorders,[15] with those from structured monitoring. The "opportunistic" records datasets used in their comparison exhibited mean numbers of site visits of 4.2 (butterflies) and 4.3 (dragonflies) across years. Earlier analyses of a subset of these dragonfly data, using single species records or short lists alone, typically found the derived trends to be "too imprecise to be very useful" (van Strien *et al.*, 2010), indicating that the type (i.e. information content) of visits available within any specific dataset is also likely to be important for accurate inference. In comparison, for British and Irish plant biological record datasets between 1990 and 2010, the within-year percentage of sites with more than one visit is around 10-25% at the 1 km scale (Fig. 5)[16]. The average number of visits received by a 1 km site within a year is around 1.4 (Fig. 6; 1.2 for bryophytes and 1.5 for vascular plants).

We may reasonably ask whether there is a contradiction between the general conclusion of studies of occupancy model design that only having two visits per site frequently leads to low quality estimates (and that therefore estimates of occupancy must be even worse if this only holds for a relatively small fraction of the site × closure period combinations considered), and the conclusion of Isaac *et al.* (2014) that their best performing occupancy model was "generally robust and powerful" in the face of simulated datasets designed to mimic the size and information content of typical sets of biological records. Two possibilities, not mutually exclusive, occur to us. First, Isaac *et al.* (2014) conducted their study in a null hypothesis significance testing/power framework, whereas studies of occupancy estimator quality have normally focused on the bias and variance in the estimation of known parameters used to simulate datasets. It is possible that occupancy models consistently produce noisy occupancy estimates with sparse opportunistic data, but still correctly estimate the linear slope of a time trend as greater than or less than zero at some percentile

---

[14] This analysis excluded rare or cryptic species that "require targeted search efforts" *a priori*, thus reducing or eliminating potential conflicts between the model of recording implied by the detection model and reality (van Strien *et al.*, 2013).

[15] Divided into single species records, short lists of two to three species, and "comprehensive" day lists (effectively any visit reporting > 3 species).

[16] A cursory look at the types of 1 km sites with repeat visits within years suggests that they can be divided into at least four cases: "honeypot" sites (e.g. nature reserves); sites very close to recorders' homes; sites that have been subject to professional surveys (e.g. repeated quadrat surveys of Sites of Special Scientific Interest); and data entry errors. A more thorough examination of the data could reveal other types. It is not necessarily the case that multiple site visits of these types actually provide the expected information on a species' detectability; for example, a recorder re-visiting a home site during a year is likely to only report new species encountered (indeed, some recorders in well-recorded areas extend this approach, i.e. only adding new species for a square within some time period rather than fully resurveying re-visited sites). Another example of revisits not providing the expected information is the case of honeypots, where the visiting recorders may repeatedly record the same population of a rare species. Such biases are likely to be more important for static plants compared to mobile animals.

(although see the following two papers: van Strien *et al.*, 2010; Erickson *et al.*, 2017). The exact impacts of this on the resulting inferences from the occupancy modelling of opportunistic biological records have not, to our knowledge, been clearly investigated in a generally applicable way: rejecting a null hypothesis of no change is not necessarily the same as accurately representing the actual changes in a species' occupancy over time, as is normally required for species indicators (e.g. Normander *et al.*, 2012).
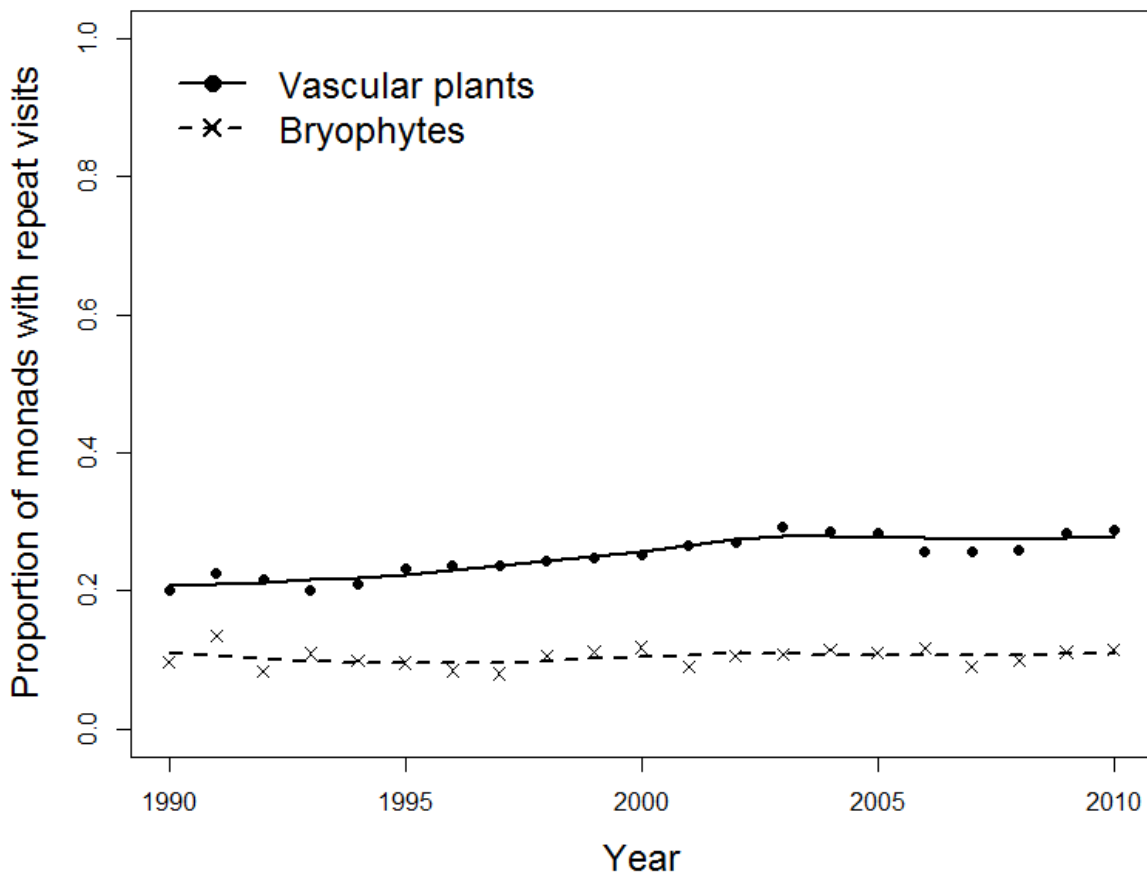
Figure 5. The proportion of monad sites (i.e. 1 × 1 km grid squares) with repeat day visits within years, for 1990-2010. The bryophyte dataset is that of the British Bryological Society (Blockeel *et al.*, 2014); vascular plant data were extracted from the BSBI Distribution Database in May 2019. 1990-2010 was chosen to match the datasets examined in Isaac et al. (2014). The averages across all years are 25.0% (vascular plants) and 10.4% (bryophytes). The lines are loess smoothers.
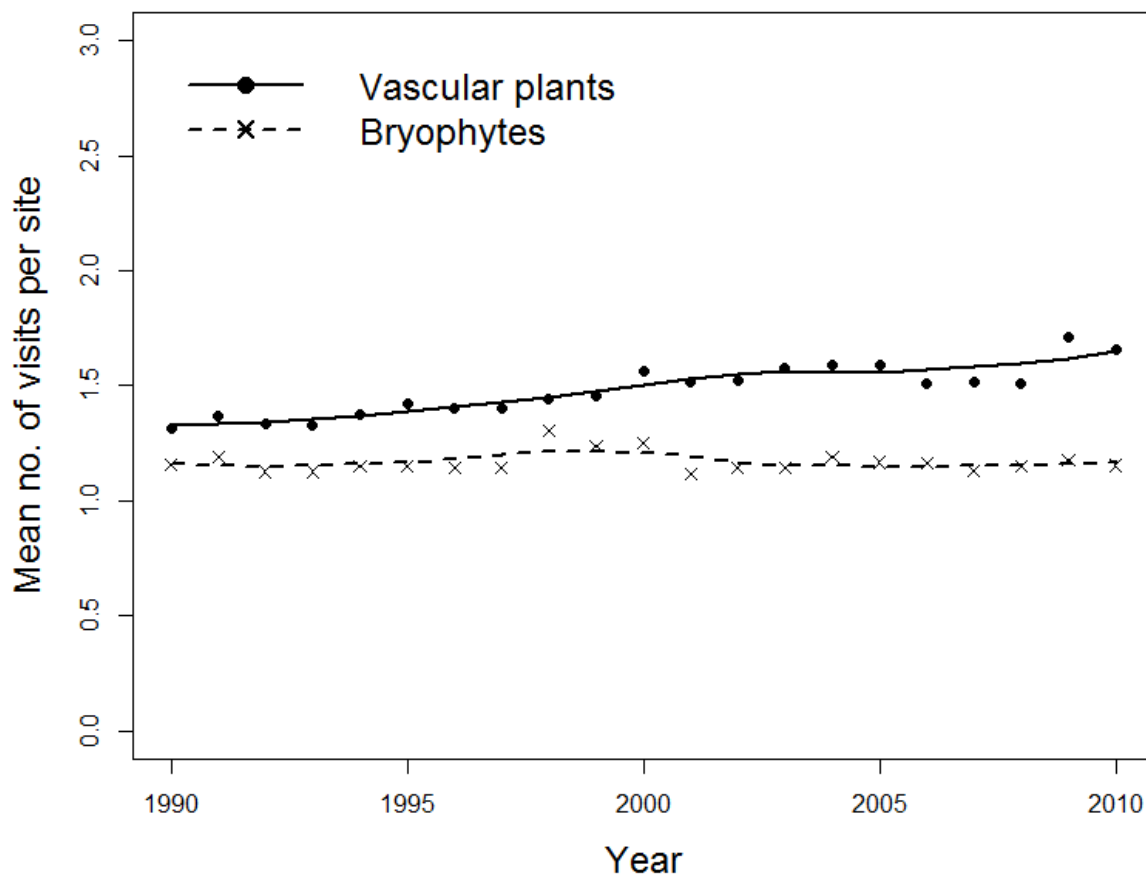
Figure 6. The mean number of visits to monad sites (i.e. 1 × 1 km grid squares) for the years 1990–2010. The bryophyte dataset is that of the British Bryological Society (Blockeel *et al.*, 2014); vascular plant data were extracted from the BSBI database in May 2019. The period 1990–2010 was chosen to match the datasets examined in Isaac *et al.* (2014). The averages across all years are 1.5 (vascular plants) and 1.2 (bryophytes). The lines are loess smoothers.

Second, Isaac *et al.* (2014) used a common pool of 1000 sites for each year of their ten year simulations. Although the within-year site visit frequencies in their simulation were parameterised using real records data, the fact that the same pool of sites was used each year must logically increase the chance that, for a given analysis, any single site will have more than one year with multiple visits relative to the real world situation. The real world consists of a much larger rotating pool of 1 × 1 km sites visited per year, at least for British and Irish plant data (e.g. Great Britain has over 300,000 land-containing 1 km squares). In existing datasets of plant biological records, although the within-year re-visited site rate averages around 18% (Fig. 5), there is relatively low overlap between the 1 km sites visited across years (Figs 7, 8); this means that, compared to the Isaac *et al.* (2014) simulation, for any given real world species there will be fewer sites where multiple years provide

271

information to a model.[17] Although some applications of occupancy modelling to British and Irish biological records (e.g. Outhwaite *et al.*, 2019) have used site filtering to only include sites with more than one year of data across the whole dataset, this does not guarantee that these sites will also have multiple visits within years, and, to our knowledge, data on patterns of re-visits in real datasets as a result of this type of filtering have not been clearly assessed in relation to statistical error (i.e. the resulting bias and variance). If filtering at the 1 km × year level results in sets of sites for a species that have very few, or zero, within-year repeat visits, then estimates of detectability may be based on little or no accurate information (MacKenzie *et al.*, 2006; van Strien *et al.*, 2010). In addition, and as already noted above in the section on reporting rate approaches, the combination of such sparseness and filtering could lead to additional biases in the dataset remaining. Appropriate checks on data availability prior to model fitting could help to identify such scenarios in real datasets, particularly if informed by more representative simulation studies and dataset-specific knowledge.

Finally, it is worth observing that occupancy modelling and Frescalo actually rely on similar logic to estimate the true underlying probability that a species is present at a site. In occupancy modelling repeated visits in time are used to estimate the probability that a species has been missed in order to adjust for this possibility; in Frescalo, an average property of the larger ecological neighbourhood of a site (Fig. 4) is used to make this adjustment, combined with the assumption that the effort-adjusted frequency of a species across a neighbourhood is approximately equal to the true rate of discovery of the species at the focal site (Hill, 2012). It could be argued that the assumptions required to construct an ecologically realistic neighbourhood for any given site are manifold and difficult to test (Isaac *et al.*, 2014), however, as noted above, the underlying assumptions of a typical occupancy modelling exercise are often equally difficult to assess, and experimental approaches to these problems have indicated that they normally do need addressing (Miller *et al.*, 2015). We feel that the comprehensive availability of floristic and land-cover data at the scale typically used for the construction of Frescalo neighbourhoods (Fig. 4), coupled with Tobler's first law of geography[18] justifying the additional use of a distance weighting, provide good support for the method of neighbourhood construction given by Hill (2012). On the other hand, the appropriateness of occupancy modelling for opportunistic data of the types overviewed here does not seem to have been fully proven, with most tests having either used relatively data-rich species, or having relied on simulations that have exaggerated the amount of information available in many opportunistic or semi-structured datasets. In addition, intrinsic biases within historic biological records datasets at the scale currently favoured by such approaches (the 1 km × year level) seem highly likely to undermine the inferential power of such models (e.g. see footnotes 7 and 8 above).

---

[17] Indeed, for the purpose of recording for the British and Irish plant atlas for 2020, volunteers have been encouraged to visit different 2 × 2 km sites within 10 × 10 km squares, not to revisit the same sites (Walker *et al.*, 2010; Groom *et al.*, 2011).

[18] "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970).

Figure 7. Jaccard similarities for bryophyte recording years based on visited monad sites (1 km grid squares) between all pairwise combinations of years, 1990–2010. As an example, a Jaccard similarity of 0.05 for a pair of years means that 5% of the total number of monads visited in two years were visited in both years, and 95% in one of the years only. The dataset is that of the British Bryological Society (Blockeel *et al.*, 2014). The lower triangle gives the Jaccard similarities; the upper triangle represents the same information as filled squares. Drawn using corrplot (Wei & Simko, 2017) in R (R Core Team, 2018).

Vascular plant monads
between-year Jaccard similarities

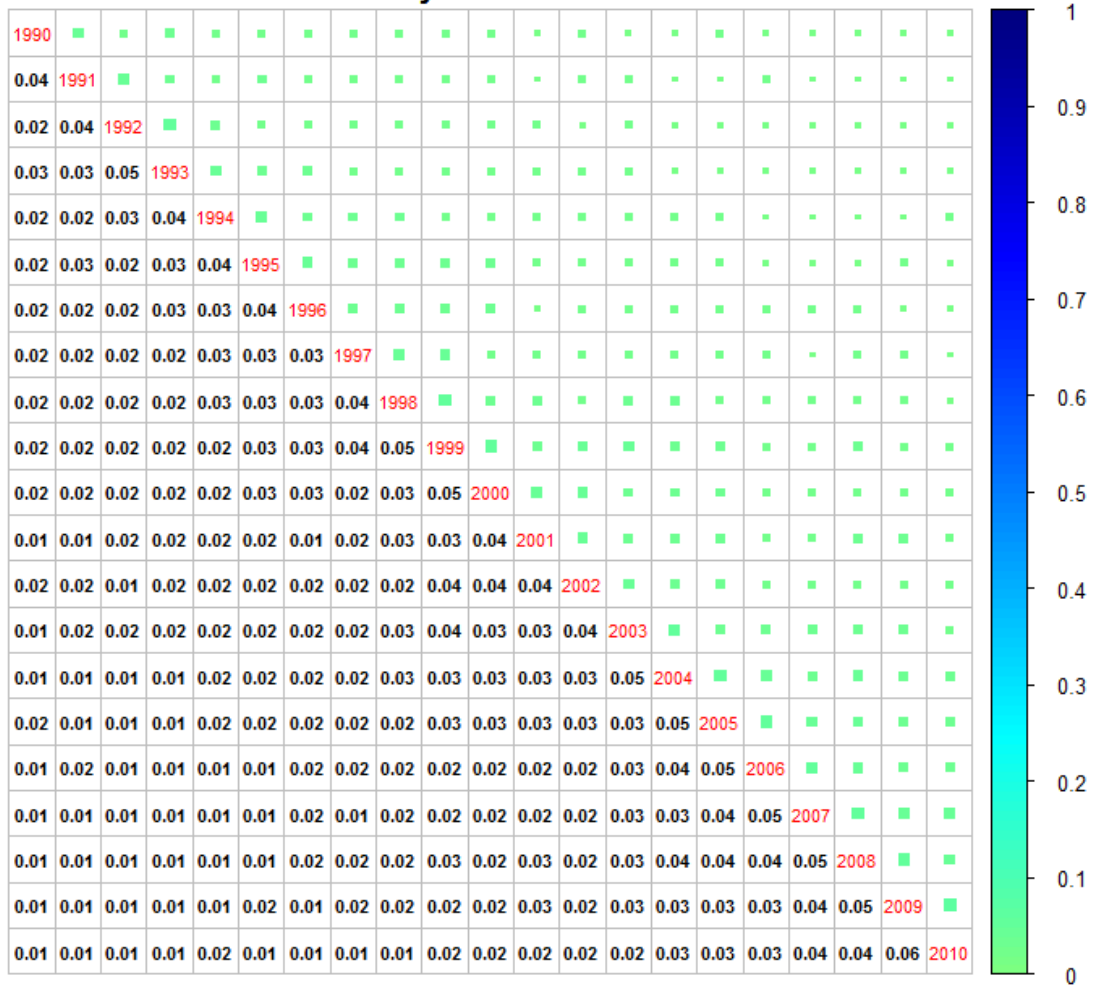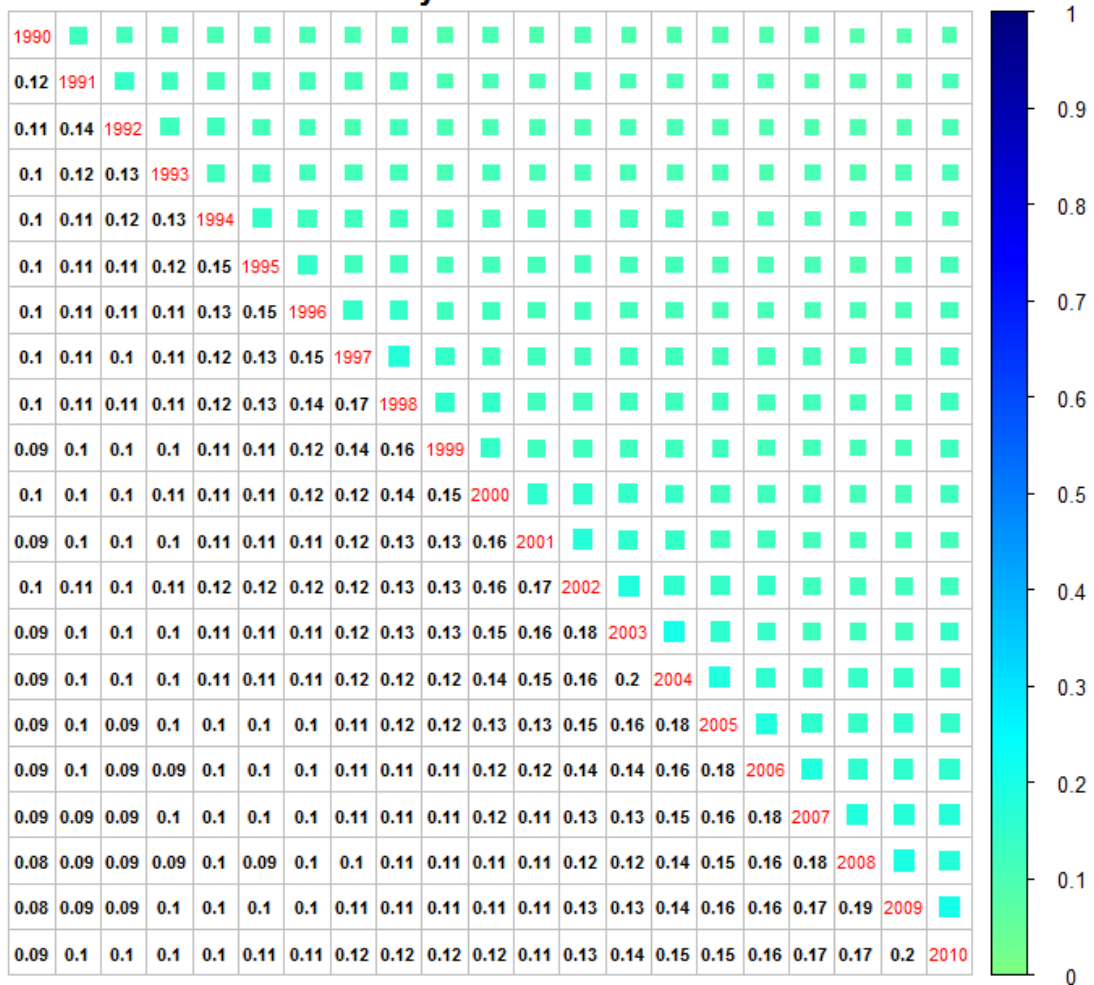|	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
1990	1990																				
1991	0.12	1991																			
1992	0.11	0.14	1992																		
1993	0.1	0.12	0.13	1993																	
1994	0.1	0.11	0.12	0.13	1994																
1995	0.1	0.11	0.11	0.12	0.15	1995															
1996	0.1	0.11	0.11	0.11	0.13	0.15	1996														
1997	0.1	0.11	0.1	0.11	0.12	0.13	0.15	1997													
1998	0.1	0.11	0.11	0.11	0.12	0.13	0.14	0.17	1998												
1999	0.09	0.1	0.1	0.1	0.11	0.11	0.12	0.14	0.16	1999											
2000	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.12	0.14	0.15	2000										
2001	0.09	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.13	0.13	0.16	2001									
2002	0.1	0.11	0.1	0.11	0.12	0.12	0.12	0.12	0.13	0.13	0.16	0.17	2002								
2003	0.09	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.13	0.13	0.15	0.16	0.18	2003							
2004	0.09	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.12	0.12	0.14	0.15	0.16	0.2	2004						
2005	0.09	0.1	0.09	0.1	0.1	0.1	0.1	0.11	0.12	0.12	0.13	0.13	0.15	0.16	0.18	2005					
2006	0.09	0.1	0.09	0.09	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.12	0.14	0.14	0.16	0.18	2006				
2007	0.09	0.09	0.09	0.1	0.1	0.1	0.1	0.11	0.11	0.11	0.12	0.11	0.13	0.13	0.15	0.16	0.18	2007			
2008	0.08	0.09	0.09	0.09	0.1	0.09	0.1	0.1	0.11	0.11	0.11	0.11	0.12	0.12	0.14	0.15	0.16	0.18	2008		
2009	0.08	0.09	0.09	0.1	0.1	0.1	0.1	0.11	0.11	0.11	0.11	0.11	0.13	0.13	0.14	0.16	0.16	0.17	0.19	2009	
2010	0.09	0.1	0.1	0.1	0.1	0.11	0.11	0.12	0.12	0.12	0.12	0.11	0.13	0.14	0.15	0.15	0.16	0.17	0.17	0.2	2010

Figure 8. Jaccard similarities for vascular plant recording years based on visited monad sites (1 km grid squares) between all pairwise combinations of years, 1990–2010. As an example, a Jaccard similarity of 0.1 for a pair of years means that 10% of the total number of monads visited in two years were visited in both years, and 90% in one of the years only. Vascular plant data were extracted from the BSBI Distribution Database in May 2019. The lower triangle gives the Jaccard similarities; the upper triangle represents the same information as filled squares. Drawn using corrplot (Wei & Simko, 2017) in R (R Core Team, 2018).

Other approaches

For the purposes of this review we have largely focused on the methods examined by Isaac *et al.* (2014), however, other methods have been proposed. One issue with some of the methods outlined above is that they have typically been implemented at the 10 × 10 km scale (i.e. 100 km^2), this has largely been because of the desire to produce trends over many decades (with much older data only being available at this resolution; Fig. 1; Rich & Karran, 2006). As any field botanist will surely appreciate, species' occupancy patterns at different spatial scales vary (e.g. Hartley & Kunin, 2003), and therefore the impression of change that we receive from any given analysis is likely to be scale-dependent (Wiens, 1989; Hodgson, 2003). In response to such considerations, Groom (2013) suggested that the well-established

spatial smoothing method known as "kriging" could be applied to plant atlas visit data collected at smaller scales (specifically the 2 × 2 km scale).[19] The presence of two systematic national surveys at this scale (Rich & Woodruff, 1990; Braithwaite, Ellis & Preston, 2006), coupled with much additional small-scale data in recent years (Fig. 1), supports this choice, and Groom took advantage of this coverage to examine changes in vascular plant "occupancy" between the periods 1978–1994 and 1995–2011. The index of occupancy used was the number of detections per number of site visits, after filtering to those visits where a minimum number of species was recorded (this cut-off was allowed to vary by region to take into account varying potential species richness across Britain; cf. Fig. 2). The kriging then calculates estimates of this index across the landscape to create a smoothed map whilst taking spatial autocorrelation into account. Groom (2013) argues that this method creates maps that are closer to the true local abundances of species at small scales.[20]

Despite the title of the paper, the method of Groom (2013) estimates the detectability of a species rather than true occupancy, thus the same criticism that was made of reporting rate methods above applies here too; that is, that the parameter being modelled is actually the confounded product of occupancy and detection (Kéry, 2011). However, given that Groom (2013) utilises filtering to only admit well-surveyed sites to his analysis, the argument is made that missing species are likely to be truly absent or rare, and therefore that the resulting smoothed detectability is a good index of true abundance across the landscape. This may not always be the case. For example, sites where rare, low abundance, species are specifically sought out during otherwise "normal" recording visits by recorders may appear to have high detectability, implying high local abundance; that is to say, detectability at the 2 × 2 km scale does not necessarily have a consistent relationship with local abundance. Rare plant species[21] should, all else being equal, have low detectability due to their low abundance in the landscape. Indeed, the correlation between fine-scale abundance (e.g. at the metre scale) and occupancy, or indices thereof, should not be pushed too far, except perhaps where the underlying statistical models have scale-invariant parameters and there is some confidence that data availability and type allow for the accurate estimation of parameters relating to observation processes (Dorazio, 2014). Some of the other motivations of Groom (2013), e.g. the ability of the method to estimate "absolute occupancy" (by which is meant a species' detectability at the 2 × 2 km scale), also apply to other methods (see footnote 20). Finally, as with other methods discussed above, the use of the approach for investigating temporal change is dependent on there being enough data across Britain and Ireland at finer scales for species' detectabilities to be estimated accurately (i.e. without bias) within time periods.

---

[19] Smith (1996) appears to have been the first to suggest this approach for interpolating between biological records data across space.

[20] Both occupancy modelling and Frescalo also output occupancy probabilities for species at whichever scale they are applied, and these can be interpreted as indices of abundance, conditional on all the assumptions that this entails (Kéry & Royle, 2016). Occupancy models can also incorporate spatial autocorrelation; Frescalo uses spatial information to estimate species' occupancy probabilities per site in the first place, although the dependence between estimates of occupancy in nearby sites is not taken into account when calculating change between time periods.

[21] We are aware of the various different definitions of rarity available to the ecologist (Rabinowitz, 1981), and are here referring to small population sizes and/or biomass relative to the area being covered by a surveyor during a recording visit.

Is there currently a clear best method for modelling species' time trends using British and Irish plant data?

In our view, the choice of an appropriate method is largely a function of choosing appropriate spatial and temporal units relative to the spatio-temporal domain for which inferences are desired. All methods require careful selection of input taxa, data processing steps, and interpretation of results for this reason (Telfer *et al.*, 2002; Rich & Karran, 2006; Pescott *et al.*, 2018). Given that the majority of older plant data is resolved only to the 10 × 10 km spatial scale, and/or to temporal units of multiple years (Fig. 1; Rich, 2006; Rich & Karran, 2006), then any researcher who wishes to model trends accurately on a multi-decadal scale will necessarily be limited in their choice of method (or be forced to ensure careful matching of data types across years).

Where simulations are concerned, the comparative performance of any set of methods will often be a function of the assumptions used to generate the fake data. For example, if hierarchical variance parameters are used in the simulation of data, then a hierarchical model should outperform a model that does not take this structure into account (e.g. Pescott *et al.*, 2016). Isaac *et al.* (2014) reported that an occupancy model with various enhancements was "generally robust and powerful" compared to other methods when presented with various types of bias typical of biological recording data. However, this conclusion may only hold when the available data have a similar structure to those simulated by these workers. When a greater proportion of sites are not revisited within years, or even across longer time-scales, and when visit-level data are simply not available (as is often the case for historic records), then other methods will likely be required to avoid complicated biases and high variance (e.g. van Strien *et al.*, 2018, 2019). In our view, further work is needed to investigate the quality of conclusions concerning temporal trends when using occupancy modelling with sparse opportunistic data. Until this understanding is developed, methods that work at broader spatio-temporal scales seem likely to deliver better results, at least for heterogeneous biological records datasets with similar properties to those discussed here.

One obvious direction for the improvement of occupancy modelling for biological records is the broadening of the temporal and spatial scales at which it is applied. This brings its own challenges, in that closure assumptions must still hold, or their potential violation be accounted for. This problem is clearest when considering broader spatial scales, when repeat visits within a closure period may be to different sub-areas of a wider area. Multi-scale and space-for-time type occupancy models may be of use here, although this is still a young field of research (Kéry & Royle, 2016 pp. 600–613). However, merely aggregating fine-scale records at a larger scale may not remove biases present at the finer scale. For example, if rarer species were recorded preferentially at a finer spatial scale in historic data, then aggregating data within a broader spatio-temporal domain will not remove these biases: the implied statistical target populations sampled over time remain unmatched.

Our review of this area leads us to expect that Frescalo is currently the preferred option for the development of time trends for British and Irish plant data; however, it should also be remembered that, for comparatively data-poor groups, methods that relax the requirement for well-estimated relative frequency curves

across all neighbourhoods of a larger territory, such as the Telfer method or variations on the reporting rate approach, may be preferred (e.g. van Strien *et al.*, 2018, 2019). Finally, and especially when faced with data that are highly likely to be biased in relation to one's aims, multiple lines of reasoning should be sought; there is no requirement for a single method, approach, or model formulation to be relied upon exclusively, particularly given that the "true" data generating process is always unknown (Steegen *et al.*, 2016; Amrhein *et al.*, 2019b).

Acknowledgements

References

Allen, D.E. 1976. *The Naturalist in Britain: A Social History*. London: Allen Lane.

Altwegg, R. & Nichols, J.D. 2019. Occupancy models for citizen-science data. *Methods in Ecology and Evolution* 10: 8–21. https://doi.org/10.1111/2041-210X.13090

Amrhein, V., Greenland, S. & McShane, B. 2019a. Scientists rise up against statistical significance. *Nature* 567: 305. https://doi.org/10.1038/d41586-019-00857-9

Amrhein, V., Trafimow, D. & Greenland, S. 2019b. Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication. *The American Statistician* 73: 262–270. https://doi.org/10.1080/00031305.2018.1543137

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T. & Jepson, P. 2015. Emerging technologies for biological recording. *Biological Journal of the Linnean Society* 115: 731–749. https://doi.org/10.1111/bij.12534

Ball, S. 2010. *Extracting trends from general biological record data with examples from the Hoverfly Recording Scheme*. Peterborough, UK: Unpublished report to JNCC.

Ball, S., Morris R., Rotheray G. & Watt K. 2011. *Atlas of the Hoverflies of Great Britain (Diptera, Syrphidae).* Wallingford: Centre for Ecology & Hydrology.

Balmer, D.E., Gillings, S. & Caffrey, B., eds. 2013. *Bird Atlas 2007-11: The Breeding and Wintering Birds of Britain and Ireland*. Thetford: BTO Books.

Blockeel, T.L., Bosanquet, S.D.S., Hill, M.O. & Preston, C.D., eds. 2014. *Atlas of British and Irish Bryophytes*. Newbury: Pisces Publications.

Boersch-Supan, P.H., Trask, A.E. & Baillie, S.R. 2019. Robustness of simple avian population trend models for semi-structured citizen science data is species-dependent. *Biological Conservation* 240: 108286. https://doi.org/10.1016/j.biocon.2019.108286

Braithwaite, M.E., Ellis, R.W. & Preston, C.D. 2006. *Change in the British flora 1987-2004*. London: Botanical Society of the British Isles.

Carey, P., Wallis, S., Chamberlain, P.M., Cooper, A., Emmett, B.A., Maskell, L.C. *et al.* 2008. *Countryside Survey: UK Results from 2007.* http://www.countrysidesurvey.org.uk/outputs/uk-results-2007

Desender, K. & Turin, H. 1989. Loss of habitats and changes in the composition of the ground and tiger beetle fauna in four West European countries since 1950 (Coleoptera: Carabidae, cicindelidae). *Biological Conservation* 48: 277–294. https://doi.org/10.1016/0006-3207(89)90103-1

Dorazio, R.M. 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography* 23: 1472–1484. https://doi.org/10.1111/geb.12216

Dunn, A.M. & Weston, M.A. 2008. A review of terrestrial bird atlases of the world and their application. *Emu - Austral Ornithology* 108: 42–67. https://doi.org/10.1071/MU07034

Dyer, R.J., Gillings, S., Pywell, R.F., Fox, R., Roy, D.B. & Oliver, T.H. 2017. Developing a biodiversity-based indicator for large-scale environmental assessment: a case study of proposed shale gas extraction sites in Britain. *Journal of Applied Ecology* 54: 872–882. https://doi.org/10.1111/1365-2664.12784

Erickson, R.A., Gray, B.R. & Eager, E.A. 2017. Estimating linear temporal trends from aggregated environmental monitoring data. *Ecological Indicators* 74: 62–72. https://doi.org/10.1016/j.ecolind.2016.10.036

Fox, R., Oliver, T.H., Harrower, C., Parsons, M.S., Thomas, C.D. & Roy, D.B. 2014. Long-term changes to the frequency of occurrence of British moths are consistent with opposing and synergistic effects of climate and land-use changes. *Journal of Applied Ecology* 51: 949–957.

Franklin, D.C. 1999. Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. *Biological Conservation* 90: 53–68. https://doi.org/10.1016/S0006-3207(99)00010-5

Gelman, A. & Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

Groom, Q.[J.], Walker, K.J. & McIntosh, J. 2011. *BSBI Recording the British and Irish flora 2010-2020. Annex 1: Guidance on sampling approaches*. Harrogate: BSBI.

Groom, Q.J. 2013. Estimation of vascular plant occupancy and its change using kriging. *New Journal of Botany* 3: 33–46.

Guillera-Arroita, G., Lahoz-Monfort, J.J., MacKenzie, D.I., Wintle, B.A. & McCarthy, M.A. 2014. Ignoring Imperfect Detection in Biological Surveys Is Dangerous: A Response to 'Fitting and Interpreting Occupancy Models".' *PLOS ONE* 9: e99571. https://doi.org/10.1371/journal.pone.0099571

Hartley, S. & Kunin, W.E. 2003. Scale Dependency of Rarity, Extinction Risk, and Conservation Priority. *Conservation Biology* 17: 1559–1570. https://doi.org/10.1111/j.1523-1739.2003.00015.x

Hill, M.[O.] & Preston, C.[D.] 2014. Changes in distribution and abundance, 1960–2013. In: *Atlas of British and Irish Bryophytes. Volume 1.*, eds. Blockeel, T.L., Bosanquet, S.D.S., Hill, M.O. & Preston, C.D. pp. 34–49. Newbury: Pisces Publications.

Hill, M.O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54: 427–432.

Hill, M.O. 2012. Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution* 3: 195–205.

Hill, M.O. & Preston, C.D. 2015. Disappearance of boreal plants in southern Britain: habitat loss or climate change? *Biological Journal of the Linnean Society* 115: 598–610. https://doi.org/10.1111/bij.12500

Hodgson, J.[G.] 2003. Change the change index? *BSBI News* 93: 44–47.

Isaac, N.J.B. & Pocock, M.J.O. 2015. Bias and information in biological records. *Biological Journal of the Linnean Society* 115: 522–531. https://doi.org/10.1111/bij.12532

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 in Ecology and Evolution* 5: 1052–1060. https://doi.org/10.1111/2041-210X.12254

Johnston, A., Hochachka, W.M., Strimas-Mackey, M.E., Gutierrez, V.R., Robinson, O.J., Miller, E.T. *et al.* 2019. Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. *bioRxiv*, 574392. https://doi.org/10.1101/574392

Jost, L. 2006. Entropy and diversity. *Oikos* 113: 363–375.

Kéry, M. 2011. Towards the modelling of true species distributions. *Journal of Biogeography* 38: 617–618.

Kéry, M. & Royle, J.A. 2016. *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS. Volume 1: Prelude and Static Models*. London, UK: Academic Press.

Kéry, M., Royle, J.A., Schmid, H., Schaub, M., Volet, B., Häfliger, G. *et al.* (2010). Site-Occupancy Distribution Modeling to Correct Population-Trend Estimates Derived from Opportunistic Observations. *Conservation Biology* 24: 1388–1397. https://doi.org/10.1111/j.1523-1739.2010.01479.x

Kirby, K.J., Bines, T., Burn, A., Mackintosh, J., Pitkin, P. & Smith, I. 1986. Seasonal and Observer Differences in Vascular Plant Records from British Woodlands. *Journal of Ecology* 74: 123–131. https://doi.org/10.2307/2260353

Latour, J. & van Swaay, C. 1992. Dagvlinders als indicatoren voor de regionale milieukwaliteit. *De Levende Natuur* 93: 19–22.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L. & Hines, J.E. 2006. *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Burlington, MA: Academic Press.

Maes, D., Isaac, N.J.B., Harrower, C.A., Collen, B., van Strien, A.J. & Roy, D.B. 2015. The use of opportunistic data for IUCN Red List assessments. *Biological Journal of the Linnean Society* 115: 690–706. https://doi.org/10.1111/bij.12530

McGill, B.J. 2014. Detection probabilities, statistical machismo, and estimator theory. *Dynamic Ecology*. https://dynamicecology.wordpress.com/2014/09/15/detection-probabilities-statistical-machismo-and-estimator-theory/

McShane, B.B., Gal, D., Gelman, A., Robert, C. & Tackett, J.L. 2019. Abandon Statistical Significance. *The American Statistician* 73: 235–245. https://doi.org/10.1080/00031305.2018.1527253

Miller, D.A.W., Bailey, L.L., Grant, E.H.C., McClintock, B.T., Weir, L.A. & Simons, T.R. 2015. Performance of species occurrence estimators when basic assumptions are not met: a test using field data where true occupancy status is known. *Methods in Ecology and Evolution* 6: 557–565. https://doi.org/10.1111/2041-210X.12342

Normander, B., Levin, G., Auvinen, A.-P., Bratli, H., Stabbetorp, O., Hedblom, M., *et al.* 2012. Indicator framework for measuring quantity and quality of biodiversity—Exemplified in the Nordic countries. *Ecological Indicators* 13: 104–116. https://doi.org/10.1016/j.ecolind.2011.05.017

Outhwaite, C.L., Powney, G.D., August, T.A., Chandler, R.E., Rorke, S., Pescott, O.L. *et al.* 2019. Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970–2015. *Scientific Data* 6: 1–12. https://doi.org/10.1038/s41597-019-0269-1

Perring, F. & Walters, S.M., eds. 1962. *Atlas of the British Flora*. London: Thomas Nelson & Sons.

Pescott, O.L., Humphrey, T.A. & Walker, K.J. 2018. *A short guide to using British and Irish plant occurrence data for research.* Wallingford, UK: NERC Centre for Ecology & Hydrology/BSBI. DOI:10.13140/RG.2.2.33746.86720

Pescott, O.L., Jitlal, M., Smart, S.M., Walker, K.J., Roy, D.B. & Freeman, S.N. 2016. A comparison of models for interval-censored plant cover data, with applications to monitoring schemes. *PeerJ Preprints* 4: e2532v1. https://doi.org/10.7287/peerj.preprints.2532v1

Pescott, O.L., Simkin, J.M., August, T.A., Randle, Z., Dore, A.J. & Botham, M.S. 2015a. Air pollution and its effects on lichens, bryophytes, and lichen-feeding Lepidoptera: review and evidence from biological records. *Biological Journal of the Linnean Society* 115: 611–635. https://doi.org/10.1111/bij.12541

Pescott, O.L., Walker, K.J., Pocock, M.J.O., Jitlal, M., Outhwaite, C.L., Cheffings, C.M., *et al.* 2015b. Ecological monitoring with citizen science: the design and implementation of schemes for recording plants in Britain and Ireland. *Biological Journal of the Linnean Society* 115: 505–521. https://doi.org/10.1111/bij.12581

Pocock, M.J.O., Roy, H.E., Preston, C.D. & Roy, D.B. 2015. The Biological Records Centre: a pioneer of citizen science. *Biological Journal of the Linnean Society* 115: 475–493. https://doi.org/10.1111/bij.12548

Preston, C.D. 2013. Following the BSBI's lead: the influence of the Atlas of the British flora, 1962-2012. *New Journal of Botany* 3: 2–14.

Preston, C.D. 2014. Recording bryophytes in Britain and Ireland, 1990–2013. In: *Atlas of British and Irish Bryophytes. Volume 1*, eds. Blockeel, T.L., Bosanquet, S.D.S., Hill, M.O. & Preston, C.D. pp. 2–11. Newbury: Pisces Publications.

Preston, C.D., Croft, J.M. & Pearman, D.A. 2002a. The 1962 *Atlas* and later recording projects. In: *New Atlas of the British and Irish Flora*., eds. Preston, C.D., Pearman, D.A. Dines, T.D. pp. 3–8. Oxford, England: Oxford University Press.

Preston, C.D. & Hill, M.O. 2019 *Cambridgeshire's Mosses & Liverworts*. Newbury, UK: Pisces Publications.

Preston, C.D., Pearman, D.A. & Dines, T.D. eds. 2002b. *New Atlas of the British and Irish Flora*. Oxford, England: Oxford University Press.

Preston, C.D. & Rorke, S. 2014. Spatial variation in recording, 1950–2013. In: *Atlas of British and Irish Bryophytes. Volume 1*, eds. Blockeel, T.L., Bosanquet, S.D.S., Hill, M.O. & Preston, C.D. pp. 12–25. Newbury: Pisces Publications.

R Core Team. 2018. *R: A language and environment for statistical computing*. Austria, Vienna: R Foundation for Statistical Computing.

Rabinowitz, D. 1981. Seven forms of rarity. In: *The biological aspects of rare plant conservation*, ed. H. Synge, pp. 205–217. Chichester, UK: John Wiley and Sons.

Rich, T. 1998. Squaring the Circles – bias in distribution maps. *British Wildlife* 9: 213–219.

Rich, T.C. 2006. Floristic changes in vascular plants in the British Isles: geographical and temporal variation in botanical activity 1836–1988. *Botanical journal of the Linnean Society* 152: 303–330.

Rich, T.C.G. & Karran, A.B. 2006. Floristic changes in the British Isles: comparison of techniques for assessing changes in frequency of plants with time. *Botanical Journal of the Linnean Society* 152: 279–301. https://doi.org/10.1111/j.1095-8339.2006.00574.x

Rich, T.C.G. & Woodruff, E.R. 1990. *BSBI Monitoring Scheme 1987–88: Volume 1 Text*. Peterborough: Nature Conservancy Council.

Robertson, M.P., Cumming, G.S. & Erasmus, B.F.N. 2010. Getting the most out of atlas data. *Diversity and Distributions* 16: 363–375.

Roy, H.E., Adriaens, T., Isaac, N.J., Kenis, M., Onkelinx, T., Martin G.S. *et al.* 2012. Invasive alien predator causes rapid declines of native European ladybirds. *Diversity and Distributions* 18: 717–725.

Royle, J.A. 2006. Site occupancy models with heterogeneous detection probabilities. *Biometrics* 62: 97–102.

Seaward, M.R.D. 1998. Time-space analyses of the British lichen flora, with particular reference to air quality surveys. *Folia Cryptogamica Estonica* 32: 85–96.

Smith, P.A. 1996. Kriging for analyses of presence/absence data in biological atlases. *Aspects of Applied Biology* 46: 151–154.

Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. 2016. Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science* 11: 702–712. https://doi.org/10.1177/1745691616658637

van Strien, A.J., Boomsluiter, M., Noordeloos, M.E., Verweij, R.J.T. & Kuyper, T.W. 2018. Woodland ectomycorrhizal fungi benefit from large-scale reduction in nitrogen deposition in the Netherlands. *Journal of Applied Ecology* 55: 290–298. https://doi.org/10.1111/1365-2664.12944

van Strien, A.J., van Swaay, C.A.M., van Strien-van Liempt, W.T.F.H., Poot, M.J.M. & WallisDeVries, M.F. 2019. Over a century of data reveal more than 80% decline in butterflies in the Netherlands. *Biological Conservation* 234: 116–122. https://doi.org/10.1016/j.biocon.2019.03.023

van Strie,n A.J., van Swaay, C.A.M. & Termaat, T. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology* 50: 1450–1458. https://doi.org/10.1111/1365-2664.12158

van Strien, A.J., Termaat, T., Groenendijk, D., Mensing, V. & Kéry, M. 2010. Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists. *Basic and Applied Ecology* 11: 495–503.

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. *Ecological Applications* 20: 2157–2169. https://doi.org/10.1890/09-0877.1

Telfer, M.G., Preston, C.D. & Rothery, P. 2002. A general method for measuring relative change in range size from biological atlas data. *Biological Conservation* 107: 99–109.

Tobler, W.R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography* 46: 234–240.

Walker, K.J., Pearman, D., Ellis, B., McIntosh, J. & Lockton, A. 2010. *Recording the British and Irish flora 2010-2020*. BSBI, London.

Wei, T. & Simko, V. 2017. *R package "corrplot": Visualization of a Correlation Matrix (Version 0.84).* Available from https://github.com/taiyun/corrplot

Welsh, A.H., Lindenmayer, D.B. & Donnelly C.F. 2013. Fitting and Interpreting Occupancy Models. *PLOS ONE* 8: e52015. https://doi.org/10.1371/journal.pone.0052015

Welsh, A.H., Lindenmayer, D.B. & Donnelly, C.F. 2015. Adjusting for One Issue while Ignoring Others Can Make Things Worse. *PLOS ONE* 10: e0120817. https://doi.org/10.1371/journal.pone.0120817

White, H.J., Gaul, W., Sadykova, D., León-Sánchez, L., Caplat, P., Emmerson, M.C., *et al.* 2019. Land cover drives large scale productivity-diversity relationships in Irish vascular plants. *PeerJ* 7: e7035. https://doi.org/10.7717/peerj.7035

Wiens, J.A. 1989. Spatial Scaling in Ecology. *Functional Ecology* 3: 385–397. https://doi.org/10.2307/2389612

Woodcock, B.A., Harrower, C., Redhead, J., Edwards, M., Vanbergen, A.J., Heard, M.S., *et al.* 2014. National patterns of functional diversity and redundancy in predatory ground beetles and bees associated with key UK arable crops. *Journal of Applied Ecology* 51: 142–151. https://doi.org/10.1111/1365-2664.12171