



**JNCC Report 785**

**Spatial and temporal representativeness of species occurrence data collected  
between 1970 and 2015 in the United Kingdom**

**Robin Boyd, Nick Isaac, David Allen and Michael Pocock**

**October 2024**

**© JNCC, Peterborough 2024**

**ISSN 0963 8091**

JNCC's report series serves as a record of the work undertaken or commissioned by JNCC. The series also helps us to share, and promote the use of, our work and to develop future collaborations.

**For further information please contact:**

JNCC, Quay House, 2 East Station Road, Fletton Quays, Peterborough PE2 8YY.

<https://jncc.gov.uk/>

[Communications@jncc.gov.uk](mailto:Communications@jncc.gov.uk)

**This document should be cited as:**

Boyd, R<sup>1</sup>., Isaac, N.<sup>1</sup>, Allen, D.<sup>2</sup> & Pocock, M.<sup>1</sup> 2024. Spatial and temporal representativeness of species occurrence data collected between 1970 and 2015 in the United Kingdom. *JNCC Report 785*. JNCC, Peterborough, ISSN 0963-8091.

<https://hub.jncc.gov.uk/assets/90742444-a64f-47a5-bdd6-f2e4c91c4de2>

**Author affiliations:**

<sup>1</sup> UK Centre for Ecology and Hydrology, Wallingford

<sup>2</sup> Natural Resources Wales, Cardiff

**Acknowledgments:**

This work was supported by the Terrestrial Surveillance Development and Analysis Partnership, comprised of the UK Centre for Ecology & Hydrology, British Trust for Ornithology, and the Joint Nature Conservation Committee.



**UK Centre for  
Ecology & Hydrology**

**JNCC EQA Statement:**

This document is compliant with JNCC's Evidence Quality Assurance Policy

<https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/>

Whilst every effort is made to ensure that the information in this resource is complete, accurate and up-to-date, JNCC is not liable for any errors or omissions in the information and shall not be liable for any loss, injury or damage of any kind caused by its use. Whenever possible, JNCC will act on any inaccuracies that are brought to its attention and endeavour to correct them in subsequent versions of the resource but cannot guarantee the continued supply of the information.

This report and any accompanying material is published by JNCC under the [Open Government Licence](#) (OGLv3.0 for public sector information), unless otherwise stated. Note that some images or tables may not be copyright JNCC; please check sources for conditions of re-use.

The views and recommendations presented in this report do not necessarily reflect the views and policies of JNCC.

## Summary

- Every year, volunteer naturalists search for animals and plants across the United Kingdom (UK) and document what they see. The resultant data – that is, species occurrence records – are vast in quantity but biased towards specific locations, times and taxa. These “sampling biases” can cause problems when one intends to draw inferences about species’ distributions from the data.
- In this report, we assess spatial and temporal biases in species occurrence data for 23 taxonomic groups collected between 1970 and 2015 in the UK. We ask two questions of the data:
  - 1) are they representative of geographic space across the UK; and
  - 2) were they sampled from the same portion(s) of the UK over time?

The answer to question 1 indicates whether the data can be described as “nationally representative”, whereas the answer to question 2 indicates whether the data can inform on changes in species’ distributions over time.

- We show that, generally, the data are not representative of geographic space in the UK, and that there is temporal variation in geographic coverage. Our results also suggest that the data are more biased for some groups than others.
- The data considered here are currently used to estimate trends in species’ distributions; these trends are then used to derive national biodiversity indicators. Our results show that the resultant trends are not nationally representative. We also find some evidence that the spatial coverage of sampling has changed over time, the implications of which are less clear. We briefly review possible approaches to deal with these biases and to understand what effects they might have on trend estimates.

## Contents

1. Introduction .....	1
2. Methods .....	2
2.1. Data.....	2
2.2. Bias assessment .....	3
3. Results .....	4
3.1. Geographic representativeness .....	4
3.2. Temporal variation in geographic coverage .....	5
4. Discussion.....	8
References .....	10
Acronyms .....	12
Appendix 1 .....	13

# 1. Introduction

Species occurrence data comprise information on the four “Ws” of biological recording: What species was seen, Where, When and by Whom. In the United Kingdom (UK), volunteer naturalists have been collecting species occurrence data in their leisure time since at least the 17th century (Allen 1976). The landscape of biological recording in the UK was formalized considerably in the mid-20th century: first by the publication of the Atlas of the British Flora by the Botanical Society of the British Isles in 1962 (Preston 2013); and then by the establishment of the Biological Records Centre (BRC), a national focus for recording of terrestrial and freshwater taxa, in 1964. The BRC now works in partnership with over 80 national recording schemes and societies (NSSs). The BRC helps the NSSs by editing and publishing Atlases; digitizing, managing and archiving data; and designing/ hosting websites (Pocock *et al.* 2015). In return, the NSSs provide species occurrence data which can be used for many research applications (Powney & Isaac 2015).

Whilst clearly a valuable resource, species occurrence data should be used with care in a research setting. Much of the data were collected opportunistically by volunteers, which presents both opportunities and challenges. It was estimated recently that there were around 70,000 active recorders in the UK (Pocock *et al.* 2015); these volunteers have the capacity to collect a considerable quantity of data across large swathes of the country. However, as volunteers, they are free to decide what taxa to search for, where and when. This leads to non-random (sometimes called nonprobability) sampling along the axes of taxonomy, space and time (Robin Boyd *et al.* 2021; Pescott *et al.* 2019). Examples might include preferential sampling of accessible and attractive locations (Geldmann *et al.* 2016; Mair & Ruete, 2016), and of interesting (e.g. rare) species (Isaac & Pocock 2015). Sampling biases of these types present challenges where the aim is to draw general conclusions about species’ distributions.

Statistical inference is the process of drawing inferences about a statistical population from samples of that population. The first task, then, is to define the statistical population. In ecology, populations are usually defined in the taxonomic, spatial and temporal dimensions (e.g. spiders in the UK over the period 1970 to present day; though we note that sometimes the taxonomic dimension is less relevant than, e.g. phylogenetic or trait space). If the data are not sampled randomly from the population, in all dimensions relevant for inference, then they are likely to be biased relative to that population (sometimes called sampling biases). Statistics derived from biased samples are likely to be biased estimators of the population parameters of interest. It might be possible to mitigate these biases, but only if they are properly understood. Hence, it is useful to screen nonprobability samples of a population of interest for potential sampling biases.

Here, we screen species occurrence data provided by 23 NSSs in the UK for sampling biases. We ask two questions of the data that are relevant to the estimation of temporal trends in species’ distributions:

- 1) are they representative of the UK geographically, and
- 2) has the same portion of the UK been sampled over time?

The answer to question one indicates whether the data can be described as “nationally representative; the answer to question two indicates whether the data are informative about changes in species’ distributions over time. We do not consider taxonomic biases in this analysis; however, the NSSs are taxon-specific, so we ask questions 1 and 2 of the data provided by each scheme separately (i.e. we assess the spatio-temporal biases in the data for each taxonomic group individually). We discuss our findings in terms of the utility of species occurrence data provided by NSSs for estimating temporal trends in species’ distributions and review some options for mitigating the biases revealed in our analysis.

## 2. Methods

### 2.1. Data

We conducted our analyses on species occurrence data provided by the 23 NSSs in Table 1. We only considered the data collected between 1970 and 2015, and which are resolved to 1 km grid cells on the British Ordnance Survey grid. This is the subset of NSS data that was used by (Outhwaite *et al.* 2019) in their seminal analysis of species' distributional changes in the UK. It is worth noting that these data are slightly outdated because the schemes have received additional data since they were shared with the BRC (from both before and after 2015). The geographic metadata provided with each record is a grid reference indicating the 1 km grid cell in which it was collected. For part of our bias assessment (see nearest neighbour index below) it was necessary to convert these grid references to coordinates – we did this by assuming each record was collected at the centre of the grid cell. We degraded the data by removing any records that were duplicated in terms of taxonomic group, grid cell and year; that is to say, records of all species in a given group each year and on a given cell were considered one record. Hence, we conducted our analyses at the taxonomic group level (e.g. bees, moths, etc.). It is not possible to conduct a bias assessment of presence-only data at the species level because there is no record of sampling activity where and when the focal species was not observed. It is more reasonable to assume that the distribution of records at the group level reflects sampling effort for that group (i.e. the target group approach, Phillips *et al.* 2009).

**Table 1.** Taxonomic groups included in this analysis and the recording schemes which provide the data.

<b>Taxonomic group</b>	<b>Recording scheme</b>
Ants	Bees, Wasps and Ants Recording Society
AquaticBugs	Aquatic Heteroptera Recording Scheme
Bees	Bees, Wasps and Ants Recording Society
Bryophytes	British Bryological Society
Carabids	Ground Beetle Recording Scheme
Craneflies	Dipterists Forum, Cranefly Recording Scheme
Dragonflies	British Dragonfly Society, Dragonfly Recording Network
E&D	Dipterists Forum, Empididae, Hybotidae & Dolichopodidae Recording Scheme
Ephemeroptera	Riverfly Recording Schemes: Ephemeroptera
FungusGnats	Dipterists Forum, Fungus Gnat Recording Scheme
Gelechiids	Gelechiid Recording Scheme
Hoverflies	Dipterists Forum, Hoverfly Recording Scheme
LeafSeedBeetles	Chrysomelidae Recording Scheme
Lichens	British Lichen Society
Moths	National Moth Recording Scheme
RoveBeetles	Staphylinidae Recording Scheme
ShieldBugs	Terrestrial Heteroptera Recording Scheme - Shield bugs and allied species

Taxonomic group	Recording scheme
SoldierBeetles	Soldier Beetles, Jewel Beetles and Glow-worms Recording Scheme
Soldierflies	Soldierflies and Allies Recording Scheme
Spiders	British Arachnological Society, Spider Recording Scheme
Trichoptera	Riverfly Recording Schemes: Trichoptera
Wasps	Bees, Wasps and Ants Recording Society
Weevils	Weevil and Bark Beetle Recording Scheme

## 2.2. Bias assessment

We assessed the data for two forms of bias:

- 1) uneven sampling in geographic space; and
- 2) temporal variation in the portion of geographic space that has been sampled.

It is important to understand which portions of geographic space have been sampled to determine whether the data are “nationally representative” and, if not, whether there are any well-sampled parts of the country. It is important to understand whether the spatial distribution of sampling has changed over time because the estimation of trends in species’ distributions is contingent on consistent sampling of the same locations over time.

To assess the data for biases, we used two “heuristics”. We use the term heuristic to acknowledge that it is not possible to determine the true extent of any biases without a large probability sample for comparison. The first heuristic, called the Nearest Neighbour Index [NNI; (Clark & Evans 1954)], indicates the extent to which the data deviate from a random (i.e. representative) distribution in geographic space across the UK. We calculated the NNI by simulating data points randomly across the UK in equal density to the empirical unique monad:year data. We then divided the mean of the nearest neighbour distances of the simulated data by the mean of the nearest neighbour distances of the empirical data. This quotient falls below 1 where the data are more clustered than a random distribution,  $\sim 1$  where the data are roughly randomly distributed, and  $> 1$  where the data are more uniformly distributed than would be expected by chance. We calculated the NNI separately for each year (1970 to 2015) and presented the results as time series. This allowed us to assess temporal variation in geographic sampling biases. It is worth pointing out that some of the NSSs do not collate data collected in Northern Ireland; these data will, of course, appear less representative of the UK than those for which Northern Irish data are available. It is also important to remember that we removed duplicate records from the same taxonomic group, grid cell and year. A corollary is that the NNIs might underestimate the degree to which the data deviates from a random distribution. Unfortunately, we do not have exact coordinates associated with the data: just 1 km grid references. Recalling that we derived coordinates by assuming the data were collected at the centre of each grid cell, it would not make sense to include all duplicates whose nearest neighbour distances would be zero.

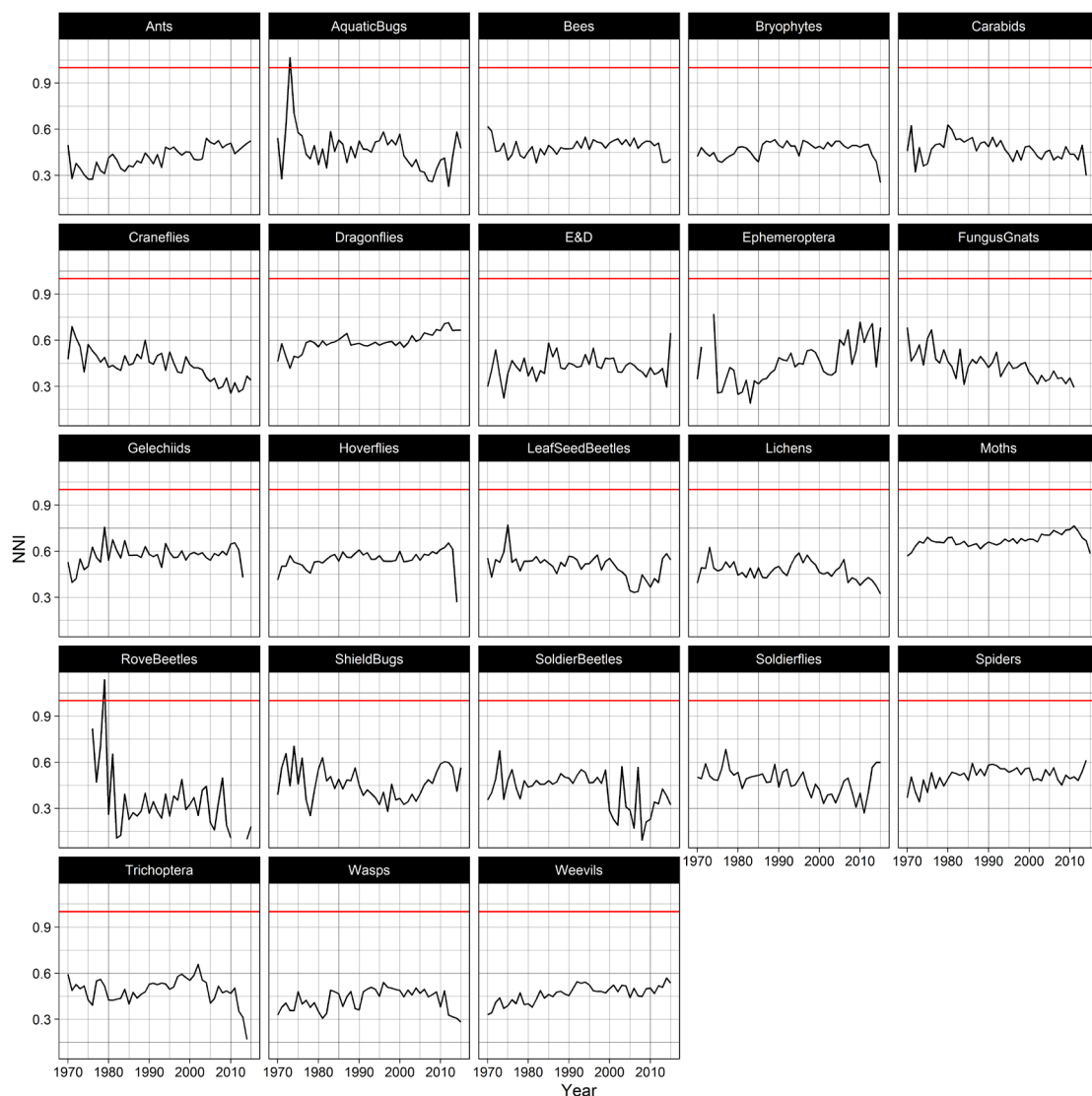
The second heuristic indicates temporal variation in geographic coverage (or spatial variation in temporal coverage) in the species occurrence data. It comprises a map showing the proportion of years in which each 1 km grid cell across the UK has been sampled. We also present this heuristic at the coarser 10 x 10 km scale for comparison in Figure 3.

We conduct our analyses using the R package *occAssess* (Boyd *et al.* 2021b).

### 3. Results

#### 3.1. Geographic representativeness

To assess the representativeness of the NSS data across geographic space in the UK, we used the NNI, which indicates the degree to which the data depart from a random (i.e. representative distribution; Figure 1). Except for rove beetles in 1976, the NNI for all groups in all years is less than one; that is to say, the data are more clustered than a random distribution. Some groups are better than others: the moths and dragonflies have relatively high NNIs in most years (usually greater than 0.6); whereas the rove beetles have low, albeit highly variable, NNIs. For some groups there appears to be a directional trend over time in the NNI. The data for ants and dragonflies tend to be more randomly distributed over time, whereas the data for crane flies become more clustered. The NNI for some groups falls drastically in the most recent years; this likely reflects a lag in data mobilisation and sharing.

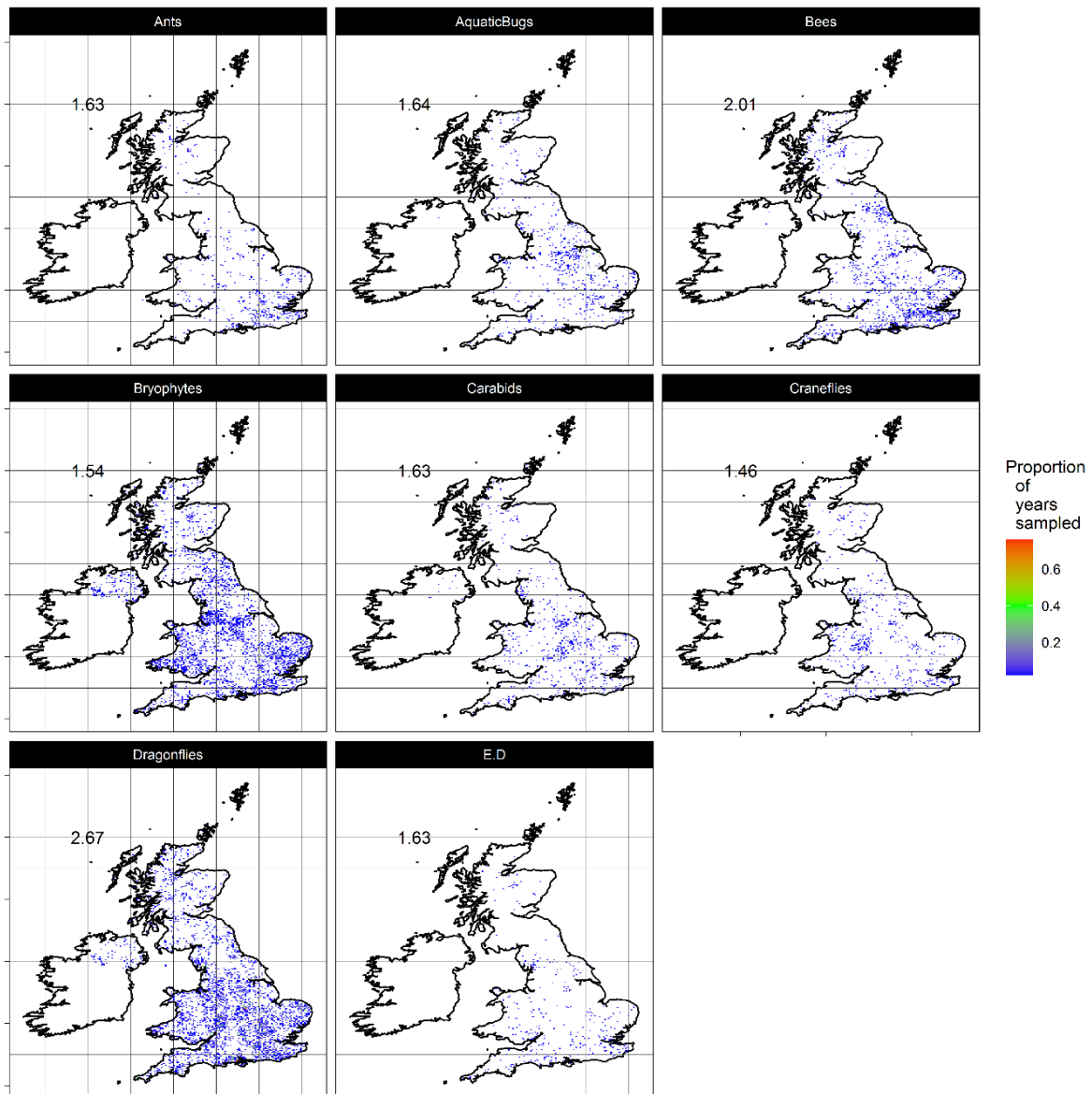


**Figure 1.** Nearest Neighbour Index (NNI) for each taxonomic group in each year. Values  $< 1$  indicate that the data are more clustered than a random distribution; values  $\sim 1$  indicate the data are roughly randomly distributed; and values  $> 1$  indicate that the data are over-dispersed relative to a random distribution. Where lines are broken there was insufficient data to calculate the NNI ( $< 10$  records).

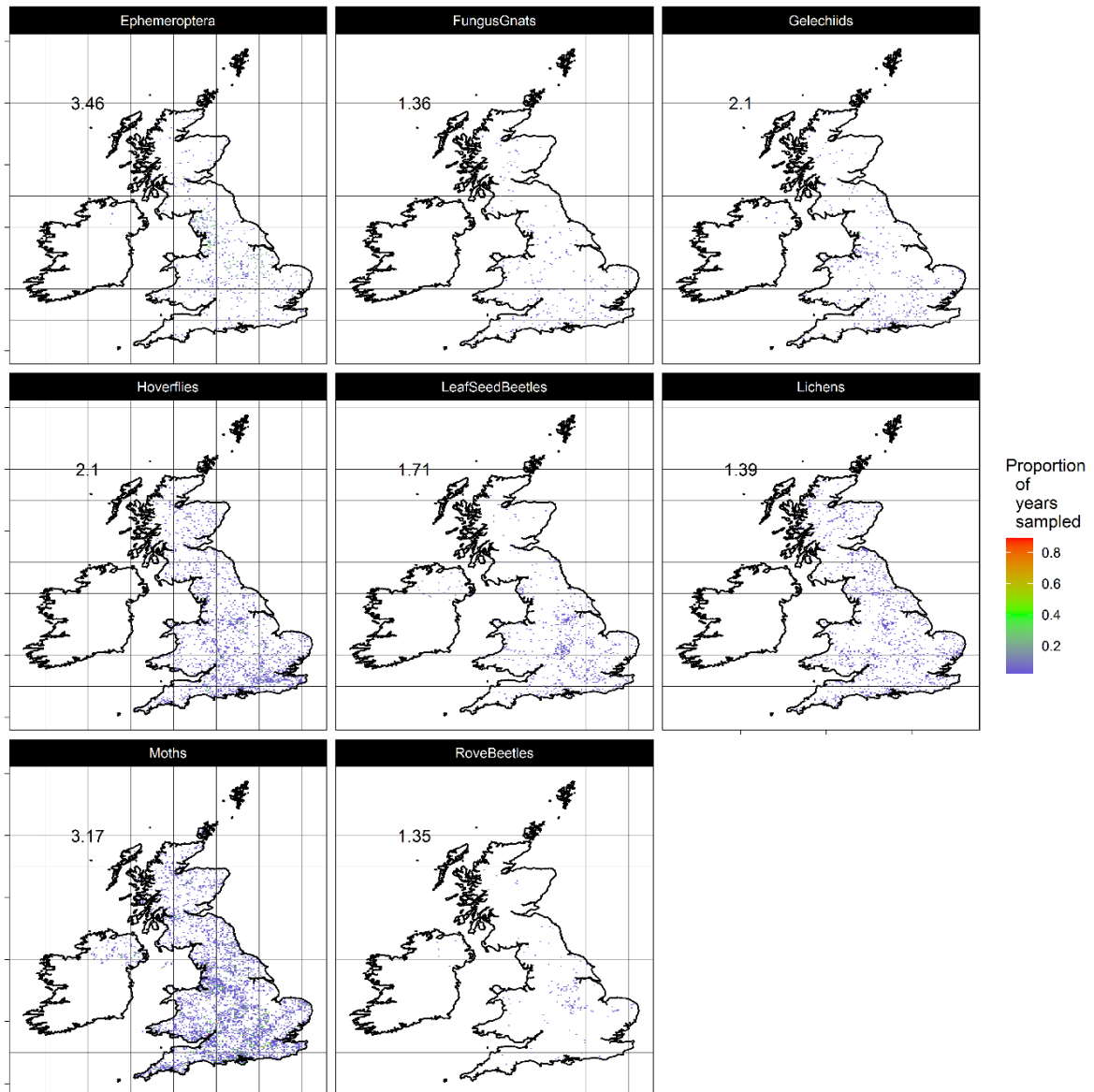


### 3.2. Temporal variation in geographic coverage

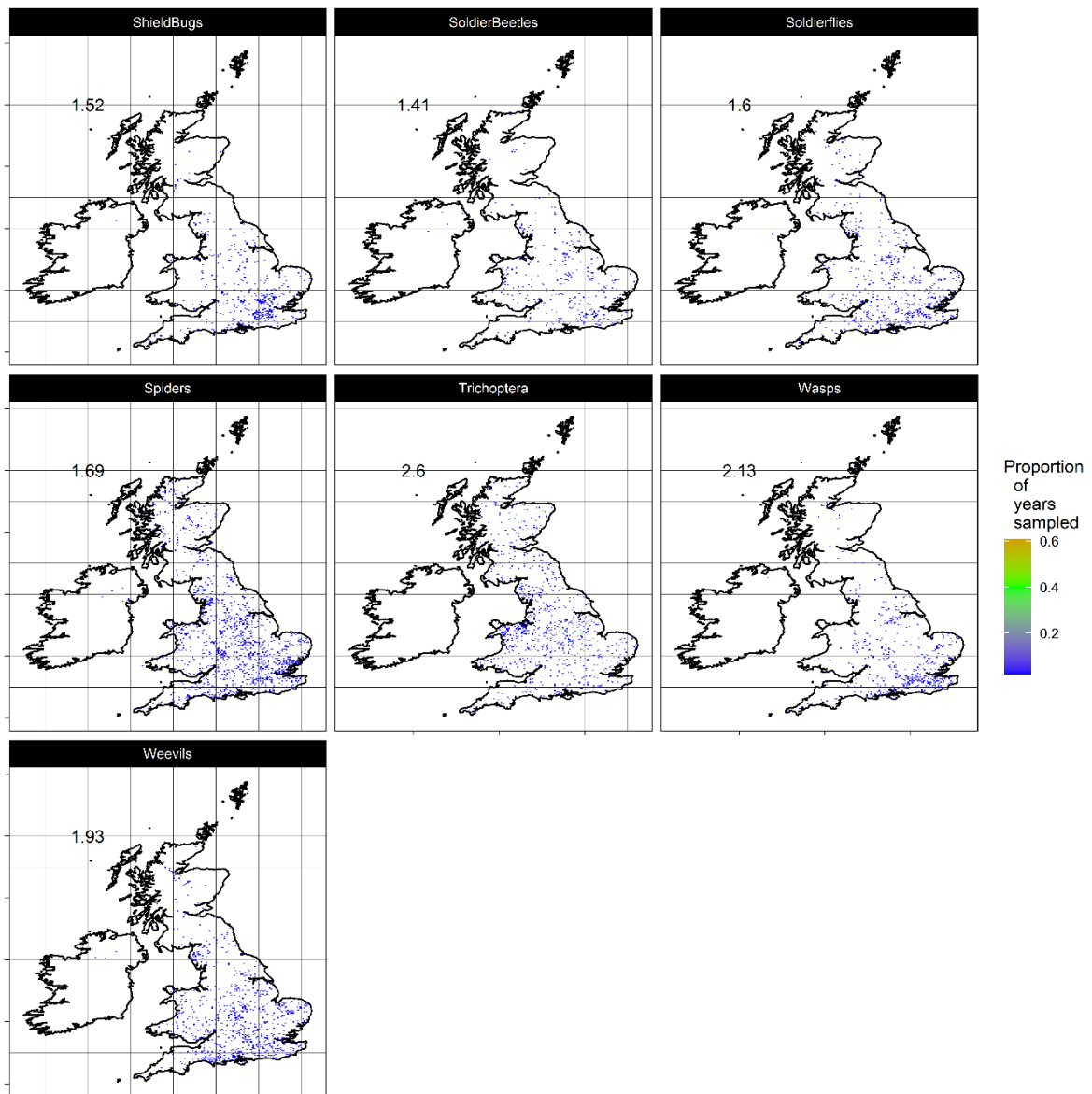
To assess the data for each taxonomic group for temporal variation in geographic coverage (or spatial variation in temporal coverage), we created a map showing, for each 1 km grid cell, the proportion of the 46 years (1970–2015) in which a record was collected (Figure 2). For all groups, the majority of grid cells were sampled in fewer than 20% of years. For some groups, a small number of grid cells were sampled in a high proportion of years. For example, records of moths and dragonflies were collected in some grid cells in almost every year. In general, a small number of grid cells are sampled in a large proportion of years, but the vast majority are only sampled in a small proportion of years. In other words, there is temporal variation in geographical coverage.



**Figure 2a.** Maps indicating the number of years in which each 1 km grid cell in the UK has been sampled for several taxonomic groups (Ants, AquaticBugs, Bees, Bryophytes, Carabids, Craneflies, Dragonflies, E&D). The numbers on each panel indicate the mean number of years in which grid cells have been sampled (excluding white cells which have never been sampled).



**Figure 2b.** Maps indicating the number of years in which each 1 km grid cell in the UK has been sampled for several taxonomic groups (Ephemeroptera, FungusGnats, Gelechiids, Hoverflies, LeafSeedBeetles, Lichens, Moths, RoveBeetles). The numbers on each panel indicate the mean proportion of years in which grid cells have been sampled (excluding white cells which grid cells have been sampled (excluding white cells which have never been sampled)).



**Figure 2c.** Maps indicating the number of years in which each 1 km grid cell in the UK has been sampled for several taxonomic groups (ShieldBugs, SoldierBeetles, Soldierflies, Spiders, Trichoptera, Wasps, Weevils). The numbers on each panel indicate the mean proportion of years in which grid cells have been sampled (excluding white cells which have never been sampled).

## 4. Discussion

In this report, we have assessed the spatio-temporal representativeness of species occurrence data collected over the period 1970 to 2015 in the UK. Specifically, we asked two questions of the data:

- 1) are they representative of geographic space in the UK, and
- 2) has the same portion of geographic space in the UK been sampled over time.

These answers to these questions are highly relevant to our ability to estimate temporal trends in species' distributions. The data were provided by 23 taxon-specific NSSs, which provided a natural way to split up the analysis: by taxonomic group. Hence, whilst we do not consider taxonomic biases explicitly, we were able to provide a taxonomic breakdown of bias severity.

The data are biased geographically, and the spatial distribution of sampling has changed over time. The NNIs for each group indicate that, in any given year, the data are clustered and not representative of geographic space in the UK (Figure 1). Moreover, the spatial distribution of sampling has changed over time. This is indicated by the fact that, among grid cells that have been sampled, the average number of years in which they have been sampled is low (Figure 2). So, it is clear the data are biased, but what impact will this have on our ability to estimate temporal trends in species' distributions?

The question of to what extent the sampling biases revealed here will affect our estimates of changes in species' distributions depends largely on whether we expect those trends to differ spatially (at some coarse scale – obviously changes in the spatial distributions at fine scale are the quantities of interest). If the trends do not differ spatially, then it does not matter which portion of the UK is sampled and when. On the other hand, if the trends do vary spatially, then spatio-temporal sampling biases have the potential to obscure our estimates. For example, a species might be faring well in one portion of the country and poorly in another; if the data were collected in the former portion of the country in one period, and the latter in the next, the one might come to the artifactual conclusion that this species is in national decline. It would be useful, then, to try and establish the extent to which trends in species' distributions vary across the UK. This might involve a shift in focus from geographic to environmental space which is more likely to explain variation in species' trends.

It would be useful to quantify exactly how much our occupancy estimates are impacted by the spatial and temporal biases revealed here; this is probably best achieved using simulated data. In 2014, Isaac *et al.* simulated species' occupancy and biased recording scenarios to test ability of various statistical models to extract robust estimates of trends in species' distributions. These simulations provided useful information on the relative abilities of various models to mitigate various types of bias. However, the simulations were not spatially or environmentally explicit; it would be useful to build on this study and test various models' abilities to extract robust trends that vary in space (environmental or geographic) in the presence of biased recording.

Earlier we stated that the biases revealed here will be most problematic if trends in species' distributions vary in space. If this turns out to be the case, then it would be prudent to take mitigating action. There are several ways we could go about this. First, we could try to correct for the biases in our models using some spatial or environmental covariates. This is an attractive and feasible option for some species but might not be possible for those with fewer records (as this would involve estimating many more parameters). Second, we could exploit additional structured data (probability samples). Some groups, such as bees and moths, are monitored in a more structured way through separate initiatives. It might be

possible to leverage these more structured data using “integrated models” (Ahmad Suhaimi *et al.* 2021; Isaac *et al.* 2020). However, the structured data tend to have more limited spatial and temporal extents than the opportunistic data so, when the two types of data are combined, it will be challenging to define an appropriate spatial and temporal domain to which the model outputs pertain. Third, we could manipulate the data. For example, we could thin the data spatially or temporally to try and obtain a more representative sample (Aiello-Lammens *et al.* 2015). Finally, we could redefine the population about which we claim to make inferences. At present we describe our estimates as “nationally representative”. It might be better to instead describe our inferences as representative of some subset of environmental or geographic space (i.e. those that have been sampled consistently over time). However, this final option might be complicated by the fact that the data pertain to different portions of time and space between taxa; this would make it difficult to define an appropriate spatial/ temporal/ environmental domain *across* taxa.

In this report we have only scratched the surface of potential biases in species occurrence data collected in the UK. In future it would be useful to conduct a more thorough assessment of the data (e.g. Boyd *et al.* 2022). Specifically, there remain three key questions that need to be answered:

- Are the data biased environmentally and taxonomically (e.g. has sampling effort shifted towards urban areas over time)? These questions could be answered through a more comprehensive report or in the form of “risk of bias” assessments (Boyd *et al.* 2021a)
- To what extent do the biases in the data bias estimates of temporal trends in species’ distributions derived from those data (this will be best achieved using simulations)?
- What are the most effective bias mitigation strategies (also best achieved using simulations)?

## References

- Ahmad Suhaimi, S.S., Blair, G.S. & Jarvis, S.G. 2021. Integrated species distribution models: A comparison of approaches under different data quality scenarios. *Divers. Distrib.* 1–10. <https://doi.org/10.1111/ddi.13255>.
- Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. & Anderson, R.P. 2015. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* (Cop.). **38**, 541–545. <https://doi.org/10.1111/ecog.01132>.
- Boyd, R., Powney, G., Burns, F., Danet, A., Duchenne, F., Grainger, M., Jarvis, S., Martin, G., Nilsen, E., Porcher, E., Stewart, G., Wilson, O. & Pescott, O. 2021a. ROBITT : a tool for assessing the risk-of-bias in studies of temporal trends in ecology. *ecoevorxiv*. <https://doi.org/10.32942/osf.io/rhvey>.
- Boyd, R., Powney, G., Carvell, C. & Pescott, O.L. 2021b. occAssess: An R package for assessing potential biases in species occurrence data. *Ecol. Evol.* <https://doi.org/10.1002/ece3.8299>.
- Boyd, Robin, Powney, G.D., Carvell, C. & Pescott, O.L. 2021. occAssess: an R package for assessing potential biases in species occurrence data. *bioRxiv*. <https://doi.org/https://doi.org/10.1101/2021.04.19.440441>.
- Boyd, R.J., Aizen, M.A., Barahona-Segovia, R., Flores-Prado, L., Fontúrbel, F., Francoy, T., Lopez-Aliste, M., Martinez, L., Morales, C., Ollerton, J., Pescott, O.L., Powney, G.D., Saraiva, A.M., Schmucki, R., Zattara, E.E. & Carvell, C. 2022. Inferring trends in pollinator distributions across the Neotropics from publicly available data remains challenging despite mobilisation efforts. *ecoevorxiv*. <https://doi.org/10.32942/osf.io/rtdyu>.
- Clark, P. & Evans, F. 1954. Distance to Nearest Neighbour as a Measure of Spatial Relationships in Populations. *Ecology* 35, 445–453. <https://doi.org/10.1007/BF02315373>.
- Geldmann, J., Heilmann-Clausen, J., Holm, T.E., Levinsky, I., Markussen, B., Olsen, K., Rahbek, C. & Tøttrup, A.P. 2016. What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. *Divers. Distrib.* 22, 1139–1149. <https://doi.org/10.1111/ddi.12477>.
- Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., Jarvis, S., Lahoz-Monfort, J., Pagel, J., Pescott, O.L., Schmucki, R., Simmonds, E.G. & O'Hara, R.B. 2020. Data Integration for Large-Scale Models of Species Distributions. *Trends Ecol. Evol.* 35, 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>.
- Isaac, N.J.B. & Pocock, M.J.O. 2015. Bias and information in biological records. *Biol. J. Linn. Soc.* 115, 522–531. <https://doi.org/10.1111/bij.12532>.
- Mair, L. & Ruete, A. 2016. Explaining Spatial Variation in the Recording Effort of Citizen Science Data across Multiple Taxa. *PLoS One* 11, e0147796. <https://doi.org/10.1371/journal.pone.0147796>.

Outhwaite, C.L., Powney, G.D., August, T.A., Chandler, R.E., Rorke, S., Pescott, O.L., Harvey, M., Roy, H.E., Fox, R., Roy, D.B., Alexander, K., Ball, S., Bantock, T., Barber, T., Beckmann, B.C., Cook, T., Flanagan, J., Fowles, A., Hammond, P., Harvey, P., Hepper, D., Hubble, D., Kramer, J., Lee, P., MacAdam, C., Morris, R., Norris, A., Palmer, S., Plant, C.W., Simkin, J., Stubbs, A., Sutton, P., Telfer, M., Wallace, I. & Isaac, N.J.B. 2019. Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970-2015. *Sci. data* 6, 259. <https://doi.org/10.1038/s41597-019-0269-1>.

Pescott, O.L., Humphrey, T.A., Stroh, P.A. & Walker, K.J. 2019. Temporal changes in distributions and the species atlas: How can British and Irish plant data shoulder the inferential burden? *Br. Irish Bot.* 1, 250–282. <https://doi.org/10.33928/bib.2019.01.250>.

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. 2009. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol. Appl.* 19, 181–197. <https://doi.org/10.1890/07-2153.1>.

Pocock, M.J.O., Roy, H.E., Preston, C.D. & Roy, D.B. 2015. The Biological Records Centre: A pioneer of citizen science. *Biol. J. Linn. Soc.* 115, 475–493. <https://doi.org/10.1111/bij.12548>.

Powney, G.D. & Isaac, N.J.B. 2015. Beyond maps: A review of the applications of biological records. *Biol. J. Linn. Soc.* 115, 532–542. <https://doi.org/10.1111/bij.12517>.

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

## Acronyms

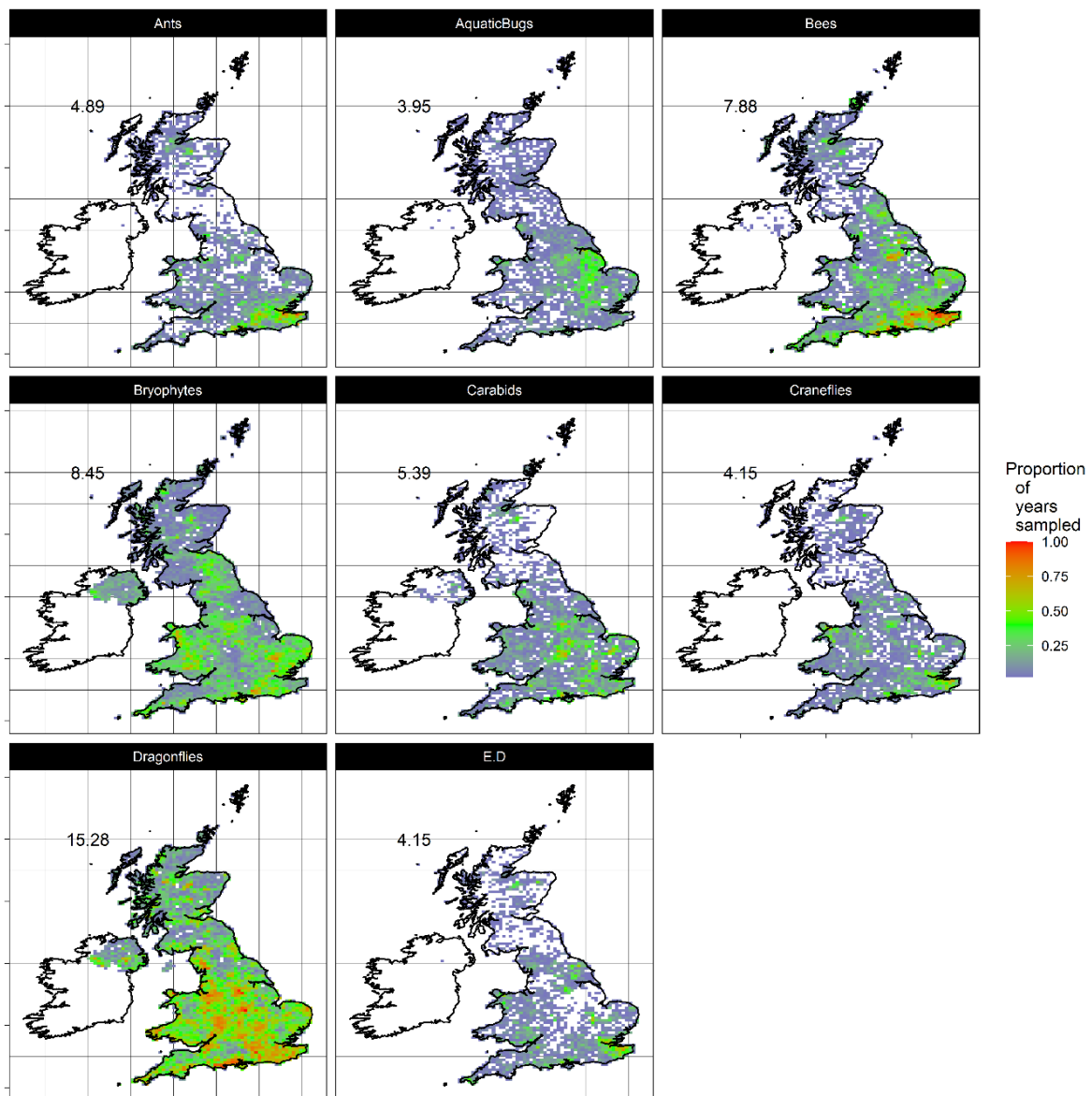
**Table 2.** Explanation of acronyms used in the report.

<b>Term</b>	<b>Definition</b>
BRC	Biological Records Centre
NSSs	National Recording Schemes and Societies
NNI	Nearest Neighbour Index

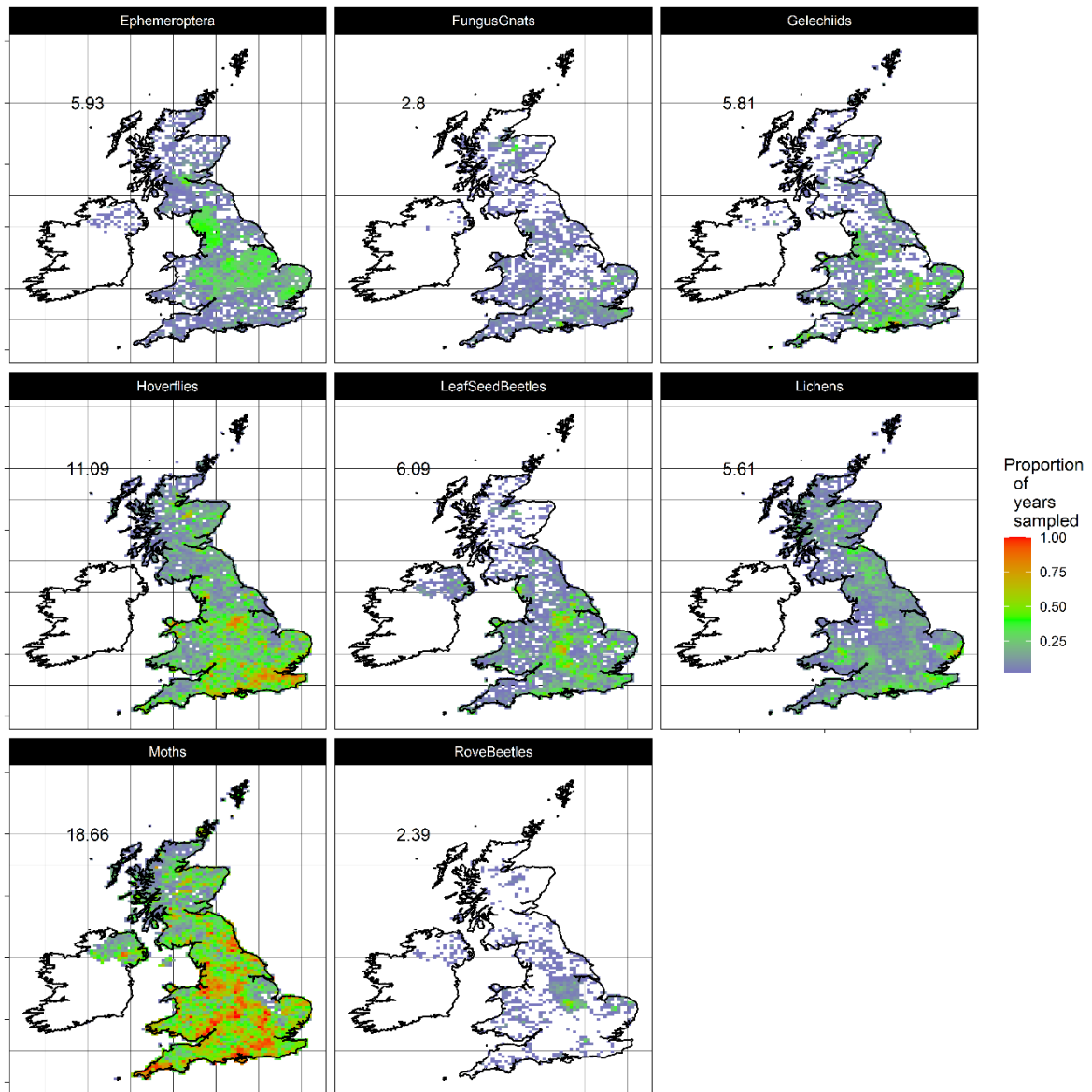


## Appendix 1

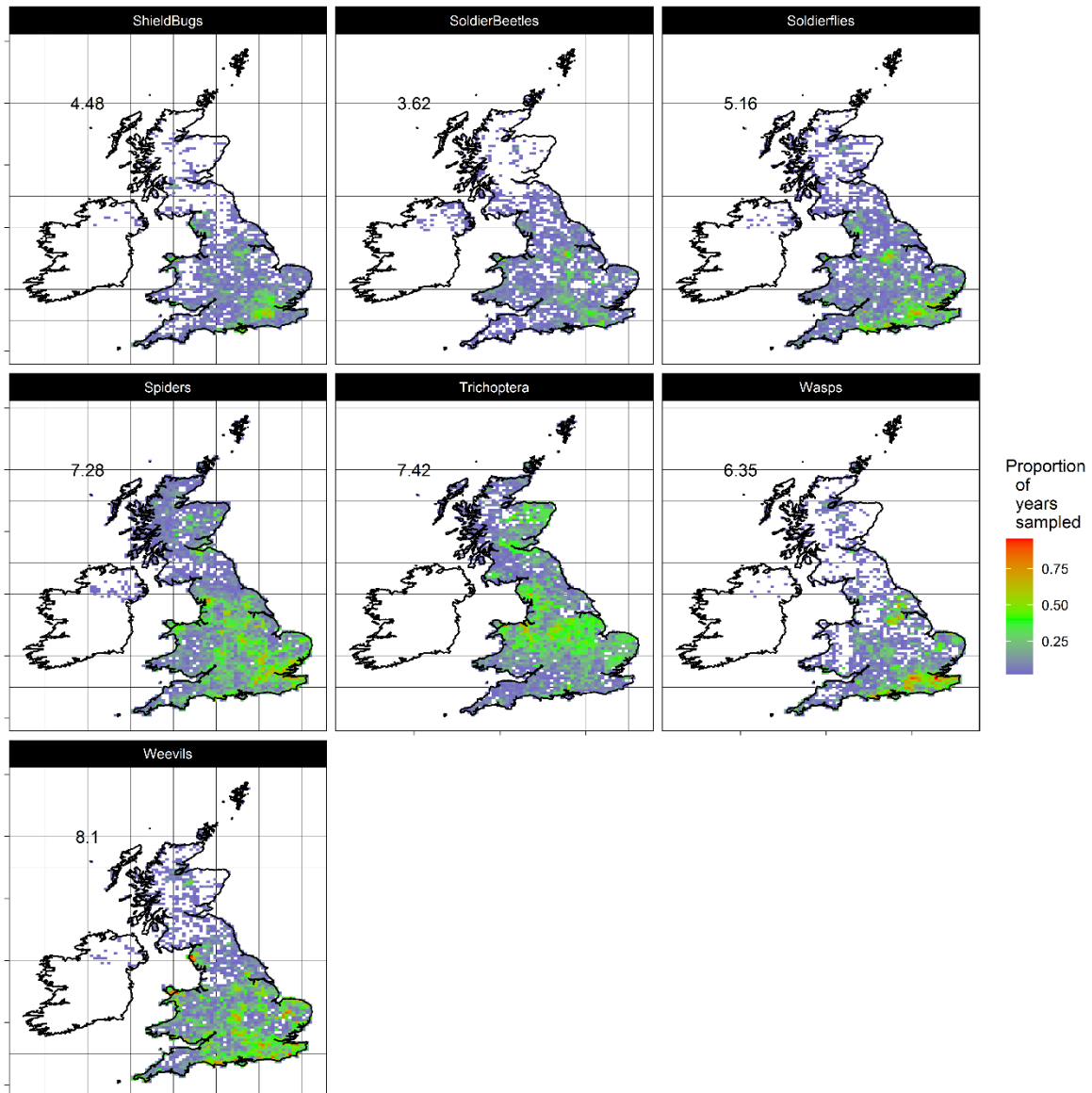
In Figure 2 we presented the number of years in which each 1 km grid cell has been sampled for each taxon group. However, as one reduces the spatial resolution of the grid cells, it is apparent that the biases are much less severe. For example, Figure 3 shows that at the 10 km resolution grid cells are generally sampled in far greater proportion of years. The flip side is that records from the same 10 x 10 km grid cell are much less likely to reflect repeat visits to the same location than multiple records from the same 1 km grid cell.



**Figure 3a.** Maps indicating the number of years in which each 10 x 10 km grid cell in the UK has been sampled for several taxonomic groups (Ants, AquaticBugs, Bees, Bryophytes, Carabids, Craneflies, Dragonflies, E&D). The numbers on each panel indicate the mean proportion of years in which grid cells have been sampled (excluding cells which have never been sampled).



**Figure 3b.** Maps indicating the number of years in which each 10 x 10 km grid cell in the UK has been sampled for several taxonomic groups (Ephemeroptera, FungusGnats, Gelechiids, Hoverflies, LeafSeedBeetles, Lichens, Moths, RoveBeetles). The numbers on each panel indicate the mean proportion of years in which grid cells have been sampled (excluding cells which have never been sampled).



**Figure 3c.** Maps indicating the number of years in which each 10 x 10 km grid cell in the UK has been sampled for several taxonomic groups (ShieldBugs, SoldierBeetles, Soldierflies, Spiders, Trichoptera, Wasps, Weevils). The numbers on each panel indicate the mean proportion of years in which grid cells have been sampled (excluding cells which have never been sampled).