Complex long-term biodiversity change among invertebrates, bryophytes and lichens

Charlotte L. Outhwaite[1, 2, 3*], Richard D. Gregory[2, 3], Richard E. Chandler[4], Ben Collen[2†] and Nick J.B. Isaac[1]

[1] *Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Wallingford, Oxfordshire, OX10 8BB, UK*

[2] *Centre for Biodiversity and Environment Research, University College London, Gower Street, London, WC1E 6BT, UK*

[3] *RSPB Centre for Conservation Science, RSPB, the Lodge, Sandy, Bedfordshire, SG19 2DL, UK*

[4] *Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, UK*

† *Deceased*

**author for correspondence: charlotte.outhwaite.14@ucl.ac.uk*

Abstract

Large-scale biodiversity change is measured largely through the responses of a few taxonomic groups. Much less is known about the trends affecting most invertebrates and other neglected taxa, and it is unclear whether well-studied taxa, such as vertebrates, reflect changes in wider biodiversity. Here, we present and analyse trends in the UK distributions of over 5,000 species of invertebrates, bryophytes and lichens, measured as changes in occupancy. Our results reveal substantial variation in both the magnitude, direction and timing of change over the last forty-five years. Just one of the four major groups analysed, terrestrial non-insect invertebrates, exhibit the declining trend reported among vertebrates and butterflies. Both terrestrial insects and bryophytes & lichens increased in average occupancy. A striking pattern is found among freshwater species, which have undergone a strong recovery since the mid-nineties following two decades of decline. We show that, while average occupancy among most groups appears to have been stable or increasing, there has been substantial change in the relative commonness and rarity of individual species, indicating considerable turnover in community composition. Additionally, large numbers of species experienced substantial declines. Our results suggest a more complex pattern of biodiversity change in the UK than previously reported.

Background

Large-scale study of the state of biodiversity is key to ensuring that conservation actions are targeted appropriately. However, high quality population time series collected on an annual basis are typically restricted to birds[1], other vertebrates[2] and butterflies[3,4]. Most invertebrates do not feature in studies of large-scale biodiversity trends[2,5] and it is unclear whether these well-studied groups reflect changes in wider biodiversity[6–8]. Invertebrates and plants constitute a substantial portion of biodiversity and support many ecosystem functions, so their lack of representation could mean that important trends are being overlooked. Recent concern about the status of under-studied groups, particularly insects[9–12] has been echoed by concerns about the quality of available evidence[13,14]. Thus, there is an urgent need to mobilise existing data and interrogate them with modern, rigorous analysis tools.

Here, we explore long-term change in UK biodiversity through an analysis of changes in the annual occupancy of numerous invertebrate groups, bryophytes and lichens. The UK is relatively well monitored in terms of national scale species' status and trends for some taxa[15,16]. The farmland bird index has declined by over 50% since 1970[17], the abundance of specialist butterflies fell by 45% from 1976-2014[18], and vascular plant communities have declined, largely in response to nitrogen deposition[19]. Conversely, bat populations in Great Britain have generally shown stable or positive trends since the late nineties[20]. The estimation of these trends has been possible through the standardised collection and analysis of abundance data. Equivalent knowledge is lacking for most biodiversity, particularly invertebrates. However, due to the rich history of species observation and recording across the UK that goes back many decades[21], extensive occurrence data are available for

these groups in the form of biological records. A record is simply a report of a species by an observer at a known time and place: most are collected opportunistically. These data have been under-utilised because of concerns over biases resulting from the unstandardized nature of data collection[22–24]. Here, we make use of recent advances in occupancy modelling techniques that address these concerns and have enabled a wider range of occurrence record datasets to be analysed[25,26]. We explore long-term change in UK biodiversity through an analysis of the average annual occupancy of invertebrate groups, bryophytes and lichens over a 45-year period. Occupancy is a measure of species range area: in this study (and following others[12,27,28]) occupancy refers to the proportion of 1km^2 grid cells in which a species is present.

We analyse outputs from national scale models of annual occupancy spanning 1970-2015 for over 5,000 terrestrial and freshwater species from 31 groups including many insect families, other invertebrate groups, bryophytes and lichens[29]. We quantify trends in average occupancy across taxa and over time and explore differences between common and rare species. We believe this represents the most comprehensive assessment of its kind, with unprecedented scale and scope for a national biodiversity assessment.

Results

Multispecies trends in occupancy

Our are comprised of occupancy estimates for 5,214 species in 31 taxonomic groups for each year from 1970-2015[29]. These estimates are derived from hierarchical Bayesian occupancy-detection models for each species and are based on 24,090,792 presence-only biological records. Multispecies trends are summarised for four aggregate taxonomic and habitat-based groups. Freshwater species (n=318) are considered separately from terrestrial species, since they are subject to different pressures. Terrestrial species are aggregated into three groups reflecting major taxonomic boundaries: terrestrial insects (n=3,089, hereafter referred to as insects), terrestrial non-insect invertebrates (mostly spiders, n=538, hereafter "invertebrates"), and bryophytes & lichens (n=1,269).

Across all 5,214 species an index of overall occupancy, estimated as the geometric mean occupancy, was 11% higher in 2015 compared with 1970 (95% credible interval: 9, 13%), contradicting the narrative that declines are pervasive. However, there were substantial differences among major groups. The most striking response was seen for the freshwater species. Although this group has experienced little net change since 1970 (+7.2%: -3.3, +19%), an increase is observed following two decades of consistent decline (Figure 1). At its lowest point, in 1994, freshwater species occupancy had declined by 47% (-51, -42%) compared with levels in 1970. Terrestrial insects show a slight increase in occupancy of 5.5% (+2.9, +8.1%) and bryophytes & lichen occupancy increased by 36% (+31, +42%). The invertebrates are the only group to experience an overall decline in mean occupancy with a reduction of 6.7% (-12, -1.6%).

Temporal variation in trends is clearly apparent when contrasting the net change in average occupancy during the first (1970-1992) and second (1993-2015) halves of the series (Figure 2). Freshwater species experienced an extreme change in trajectory, with the fastest declines exhibited across groups experienced before 1992, and the fastest increase across groups, post-1992. Terrestrial insects and invertebrates show opposite patterns with the former presenting an increase pre-1992 and a decline post-1992, and the latter declining initially then stabilising. Bryophytes & lichens respond with a slow increase pre-1992 then a more rapid increase post-1992. None of the five groups declined consistently across the time period assessed: the freshwater and insect groups experienced a reversal in average trend from the early and late parts of the time-series. This disparity across groups highlights the potential variability in response to specific drivers of change and/or response to a variety of drivers.

We detected variation in the magnitude and timing of changes in the status of rare and common species among groups (Figure 3). For both the insects and the bryophytes & lichens, the rare species showed greater changes in occurrence than the common species. Conversely, the pattern of change of the common Freshwater species indicates an earlier start to the recovery phase than for rare species. Invertebrates present no difference in response between rare and common species implying that the composition of invertebrate communities has changed less than in the other groups. Differing responses between rare and common species observed for some groups is indicative of species turnover in local communities with rare species becoming more common or vice versa.

There is considerable heterogeneity within the four groups, with some taxa facing substantial declines not apparent from the average group level change (Figure 4), this is particularly evident for the insect group. Most taxa in the freshwater group show the U-shaped trajectory of the average response. The overall decline of the invertebrates is mainly driven by spiders during the early period, and by terrestrial molluscs more recently. Bryophytes & lichens increase overall, but the timing of these increases differ. Across all 31 taxonomic groups, ten increased in mean occupancy (95% credible intervals for the year 2015 do not span 100) and five decreased (see Table S1 and Extended Data Figure 1 for more details).

Aggregating species level change into indicators of average occupancy over time hides the variation among species. Within any group there will be winners and losers whose response is more extreme than that of the average[30]. Analysis of the annual growth rates (year to year change) in occupancy of individual species reveals that although the mean change in occupancy is an increase of 11%, there are species that have undergone substantial declines (some of which started out relatively common), as well as initially rare species that have increased dramatically over time (Figure 5). There is little correlation between average occupancy and average growth rate for any of the four major groups (Pearson's correlation coefficient: Freshwater = -0.078, Insects = 0.002, Invertebrates = -0.061 and Bryophytes & lichens = -0.0001). Although most species can be found around the zero line, there are large numbers of species whose distribution changed dramatically. The lowest decile (n=529) of species' growth rates

is -2.1%, i.e. 10% (529) of species declined by at least 2.1% each year. Across the full range of 46 years in our dataset, this corresponds to a loss of 62% of previously occupied grid cells. The upper decile is 2.4%, i.e. 10% of species increased by at least this amount each year, corresponding to a tripling in distribution over 46 years. Determining drivers of change and those species most likely to exhibit strong responses will aid in mitigation of future losses.

Discussion

Our analysis of changes in occupancy for over 5,000 UK species has shown that only one of the four aggregate groups, the invertebrates, presents a decline in average occupancy. Based on widely reported declines in the abundance of UK birds[31] and butterflies[3], and for other taxa globally[2,32], one might expect to see declines in occupancy, but our results confound this expectation. Not only is the overall net change positive, the direction of change is different among groups and there is enormous variation among taxa in the temporal pattern of change and the relative fates of rare and common species. By broadening the taxonomic scope of investigation to these lesser-known groups, our results challenge the received wisdom that all biodiversity change is loss, and it is both pervasive and unalterable.

The increasing trajectories among bryophytes & lichens (since 1970) and the freshwater group (since 1994) both suggest a beneficial impact of environmental regulations and consequent management changes in reversing biodiversity declines. Bryophytes and lichens are known to be particularly sensitive to acidic pollutants such as sulphur dioxide[33], and the atmospheric concentrations of these chemicals have been declining since the first Clean Air Act of 1956. Overall, freshwater species experienced substantial declines up until the mid-nineties followed by a recovery to 1970 levels. This U-shaped trajectory is replicated for four of the six freshwater taxa, suggesting a common response to a single driver. It is notable that the lowest mean occupancy for the freshwater group follows shortly after the introduction of the European Urban Wastewater Treatment Directive in 1991, and the step change in regulation of the water industry after privatization in 1989[34]. Improvements in water quality have been linked to increases in family level richness of freshwater invertebrates in Great Britain from the early nineties onward[35,36] and historically at sampling locations in England[37]. The recent increases we report must be seen in the context of much larger declines that probably occurred over two centuries since the industrial revolution. This makes it hard to interpret the relative trajectories of rare and common species, or the apparent stasis in freshwater occupancy since 2005. Nonetheless, the recoveries in bryophytes, lichens and freshwater species, concomitant with improvements in air and water quality over recent decades, contributes to a growing trend of optimistic narratives in debates about biodiversity conservation[9,38].

Caveats about declines prior to 1970 also apply to the other major groups under consideration. It is well-established that major transformations of the UK landscape occurred during the middle parts of the 20th century[39,40] so care is needed to avoid the pitfalls of shifting baseline syndrome[41,42]. Values

seen in 1970 must not be seen as a target to be reached, since it is likely that levels in 1970 were also considerably lower than earlier in the century.

It should be noted that both recent colonist species and non-native species (such as the Harlequin ladybird) are included in the set of species analysed here, since they also contribute to UK biodiversity change[43]. However, their low numbers (56 species, just over 1% of the total - see Methods for a breakdown) result in very little influence on the multispecies trajectories presented here.

Our results demonstrate the insights that can be gathered from careful analysis of presence-only occurrence records. However, they raise questions about how trends in occupancy should be interpreted in the context of more widely-used metrics, such as changes in abundance[2], species richness[44], biomass[10,45] or other measures of range size[46,47]. For example, changes in total abundance[48], in total biomass[10] or in average species' abundance[3,31] measure similar, but subtly different, facets of biodiversity. Here, our focus is on the average change across species (as in biodiversity indicators), so we assess average occupancy across species and changes therein. In general, we would expect trends in average occupancy to under-estimate trends in average abundance, but that the two metrics would be closely correlated[49–51]. Indeed, van Strien *et al,* when assessing trends in both the occupancy and abundance of butterflies in the Netherlands observed greater changes in average abundance than in average occupancy, but both metrics present an overall negative trend[52]. Similarly, species richness and occupancy can be thought of as alternative ways of summarizing a three dimensional space-time-species data cube in which the cells of the cube represent the binary presence-absence state[53,54]. Typically, if average occupancy goes up, average richness per grid cell will also go up, so trends in richness and occupancy should be closely correlated.

This study takes us a step closer to understanding the status of UK biodiversity by exploring patterns of change for groups of species that have previously been neglected in large-scale studies of change. It is clear that occupancy is a valuable tool for assessing patterns over time when abundance data are not available. Over time, available models and required computation time will improve, enabling the wider application of occupancy modelling with models developed specifically to suit each dataset rather than the one-size-fits-all approach taken to generate the estimates used here. Although more bespoke modelling approaches could deliver greater insights when exploring change for individual species, we believe this approach is a reasonable trade-off since we are interested in aggregated trends. However, this is currently the only and best information available for most of these species groups. Importantly, this work presents just one facet in the multifaceted nature of biodiversity change[55].

Methods:

Details of the methods used in the analyses presented here are described below. These analyses can be recreated using the original data through the associated R package *UKBiodiversity*. This R package is available from GitHub (https://github.com/CharlieOuthwaite/UKBiodiversity). The package vignette contains detailed instructions for reproducing each of the statistics and figures presented here.

Data

A species occurrence dataset, as presented by Outhwaite *et al*[29], was used to assess annual species occupancy and determine estimates of annual growth rate. This dataset includes 1,000 samples from the posterior distribution of occupancy estimates for 5,293 species from 1970 to 2015. These data are freely available from the Natural Environment Research Council (NERC) Environmental Information Data Centre (EIDC, https://doi.org/10.5285/0ec7e549-57d4-4e2d-b2d3-2199e1578d84)[29]. An associated Shiny app (https://shiny-apps.ceh.ac.uk/speciesplotviewer/) can be used to view occupancy and detection plots for individual species, although we emphasise that the models developed here may not be optimal for every individual species considered: the plots should not be used uncritically for single-species assessments, therefore. These estimates are derived from occurrence records analysed using a Bayesian occupancy modelling framework based on that of Outhwaite *et al.*[25]. The model used is a hierarchical model that separates occupancy and detection to allow for the accounting of imperfect detection. Imperfect detection and other biases are common in occurrence record datasets such as those used by Outhwaite *et al*[29], however occupancy modelling has been shown to be the most appropriate method for analysing this form of data[26]. The model used to generate the occupancy estimates analysed here is made up of the following submodels.

The state model describes the true occupancy state, z_{it}, of a site i in year t (equations (1) and (2)). z_{it} will be 1 when a site is occupied and 0 if not occupied. The true occupancy, z_{it}, then takes a Bernoulli distribution:

$$z_{it} \sim \text{Bernoulli}(\psi_{it}), \quad (1)$$

The logit of the probability of occurrence, ψ_{it}, varies with both year and site:

$$\text{logit}(\psi_{it}) = \log\left(\frac{\psi_{it}}{1-\psi_{it}}\right) = b_{tr(i)} + u_i\,, \quad (2)$$

$b_{tr(i)}$ is the year effect for year t in region r in which site i is found. u_i is the site effect.

The observation model describes the data collection process. It is conditional on the true occupancy state z_{it}. p_{itv} is the probability that a species will be observed on a single visit, given the species is present at that site. The observation, y_{itv}, is then drawn from a Bernoulli distribution conditional on the true occupancy state:

$$y_{itv}|z_{it} \sim \text{Bernoulli}(p_{itv}.z_{it}) \quad (3)$$

Variation in detection probabilities p_{itv}, per visit are described as*:*

$$\text{logit}(p_{itv}) = \log\left(\frac{p_{itv}}{1-p_{itv}}\right) = a_t + \beta_1 * datatype2_{itv} + \beta_2 * datatype3_{itv}, \quad (4)$$

where β_1 and β_2 estimate differences in logit(p_{itv}) for a list length of 2-3 (datatype2) and of 4+ (datatype 3) respectively, relative to a list length of one and a_t is a year effect.

Full details of the model used can be found in the data paper associated with the species occurrence dataset[29]. All species included in the species occurrence dataset are used here except for the Rove Beetles (79 species). The Rove Beetles only have data for 1980 onwards, since the indicators we present start at 1970 it was decided that this group would be removed to maintain a simple indicator method. We therefore assess changes in occupancy over time for 5,214 species. The dataset contains derived estimates of annual occupancy, with appropriate measures of uncertainty, for species with at least 50 records (see [29] for more detail). Although this is very few records across the 45-year time period, increasing this threshold to 200 records per species did not materially change the trajectories of the aggregate group level change observed in Figure 1 (see Extended Data Figure 2 in the). A higher threshold was therefore not deemed necessary.

We conducted a posterior predictive checking exercise to assess whether our models are appropriately parameterized and not over- or under-fit (see below). Since the inferences in the paper are about trends in large multispecies taxa, we calculate summaries for groups of species rather than for individual species.

Species grouping

Species were aggregated into four major groups. Freshwater species were analysed separately, and the terrestrial species were split into three groups (Table S2). Freshwater molluscs were separated from terrestrial molluscs based on expert opinion and web-based searches. Each of the 31 taxonomic groups considered here represents either a single family (e.g. ants, family Formicidae), a selection of families (e.g. Plant Bugs) or a grouping of higher taxonomic rank (e.g. caddisflies, order Trichoptera). These groups represent the sets of species recorded by separate recording schemes within the UK (more information on these schemes can be found here: https://www.brc.ac.uk/recording-schemes).

Recent colonists and non-native species have not been excluded from this study. Due to their small number within each of the four aggregate groups their influence on the overall patterns of change in this study is minimal (Table S3). Most of the species known to be recent colonists to the UK that are included in this study are moths (n = 32). Species were identified from two lists: first, a list of recent colonists into the UK established from the literature and a list of established non-native species derived from the GB Non-native Species Secretariat Information Portal species register.

Composite trends

The composite multispecies indicators (as shown in Figures 1, 3 and 4) are calculated as the geometric mean across species[56]. To generate these indicators, we used the posterior samples of the occupancy estimates for each species in each year provided within the "POSTERIOR_SAMPLES" folder of the data source[29]. These consist of 1000 samples describing the proportion of occupied sites per species per year. For each group of species (per taxa or per aggregate group), the 1000 samples for each species within that group or taxon were assessed. For each group (or taxon), the multispecies structure was represented in the following way:

$$Y_{ist} = \alpha_t + \beta_{st} + \varepsilon_{ist} \; (t = 1, \dots, T; s = 1, , \dots, S; i = 1, \dots n) \, , \qquad (5)$$

where *T* is the total number of years considered, *S* is the total number of species, *n* (=1000 here) is the number of posterior samples and Y_{ist} is the logarithm of the *i*th posterior occupancy sample for species *s* in year *t*. In this representation, α_t denotes the posterior index of overall log-occupancy in year *t*; β_{st} is an adjustment representing the extent to which species *s* differs from this overall value; and the $\{\varepsilon_{ist}\}$ are independent random variables, each with expectation zero over the posterior distribution, representing the posterior variation for each species:year combination. To ensure that the representation (5) is unique and that α_t has the desired interpretation, the constraint $\sum_{s=1}^{S} \beta_{st} = 0$ is imposed for each year. Without this constraint there are *T*+*TS* coefficients (the $\{\alpha_t\}$ and $\{\beta_{st}\}$ in equation (5)), but only *TS* species:year combinations in the posterior dataset: with *T* constraints in total on the $\{\beta_{st}\}$, the redundancy is removed.

Define $\bar{Y}_{it} = S^{-1} \sum_{s=1}^{S} Y_{ist}$, the mean of the *i*th posterior samples across all species at time *t*. Under the assumption that the posterior occupancy probabilities for different species are independent, the quantities $\{\bar{Y}_{it} : i = 1, \dots, n\}$ are themselves samples from the posterior distribution of overall log-occupancy for year t (the independence assumption is needed to justify combining posterior samples across species). Each has expected value

$$\mathbb{E}(\bar{Y}_{it}) = S^{-1}\mathbb{E}\left(\sum_{s=1}^{S} Y_{ist}\right) = S^{-1}\mathbb{E}\left[\sum_{s=1}^{S}(\alpha_t + \beta_{st} + \varepsilon_{ist})\right] = \alpha_t \, , \qquad (6)$$

the remaining terms vanishing because $\sum_{s=1}^{S} \beta_{st} = \mathbb{E}(\varepsilon_{ist}) = 0$. Therefore, if *n* is large then the mean of the $\{\bar{Y}_{it} : i = 1, \dots, n\}$ will be close to α_t, the desired index of overall log-occupancy. Thus, $e^{\bar{Y}_{it}}$ corresponds to the geometric mean occupancy, which we present in Figure 1 and 4, after rescaling to start at 100 in 1970, with the mean and 95% quantiles taken to summarise the uncertainty of this index. To determine whether the patterns observed in Figure 1 were due to specific species responses within these groups rather than random variation, the indicators were all recalculated for alternative datasets produced by randomly reassigning species to major groups in such a way that each group contained the same number of species as in reality: this reassignment, which is carried out in the spirit of a permutation test (ref[57] Section 3.3) ensures that any inter-group differences in the alternative datasets can only be due to random variation. Results, such as those in Extended Data Figure 3, do not show the kinds of

patterns that are observed in Figure 1: these patterns therefore represent genuine inter-group differences rather than random inter-species variation.

Indices presenting changes in rarity and commonness of species over time (Figure 3) were calculated in a similar way. Specifically, rather than estimating the geometric mean occupancy $\{\bar{Y}_{it}\}$ across species posterior samples within a major group, we estimate the quantiles $\{Q_{it}\}$ corresponding to rare species (25th percentile) and common species (75th percentile). Figure 3 then summarises the posterior distributions $\{100e^{Q_{it}-Q_{i1}}\}$, for each of these quantiles.

Percentage change in occupancy of all species and aggregate groups, presented in the main text, was calculated using the geometric mean occupancies for the first, $\bar{Y}_{i1,}$ and last years, $\bar{Y}_{iT}$, (1970 and 2015 respectively):

$$\delta_{\mathrm{i}} = 100\left(e^{\bar{Y}_{iT}} - e^{\bar{Y}_{i1}}\right)/e^{\bar{Y}_{i1}}, \quad (7)$$

Similarly, annual growth rates, λ_i, for each species, provided within the data source, as:

$$\lambda_{is} = 100\left(\left(\frac{Y_{isT}}{Y_{is1}}\right)^{\frac{1}{T}} - 1\right) \quad (8)$$

For the calculation of species' growth rates, we used only the first (t=1) and last (t=T) years in which the species was recorded, in order to avoid any bias due to the potential information in the priors[25].

Sensitivity of the indicators to the number of records of a species

The dataset of annual occupancy estimates for UK species that formed the basis of this analysis included modelled outputs for species which had 50 records or more within the raw dataset (see ref[58] for more details). This is a very low number of records across the dataset that encompasses the years 1970 to 2015. To ensure that species with few records were not having a large influence on the overall trends and patterns of change over time, we recalculated the major group indicators presented in Figure 1 of the main text but used higher thresholds of the number of records of a species (Extended Data Figure 2). Overall, higher thresholds (including 75, 100, 150 and 200 records minimum per species) did not have a major effect on the patterns of change over time observed in Figure 1. Trends differed slightly, such as for the Bryophytes & lichens in some instances, but no major differences or changes in direction were seen.

Variability within major group level indicators

Considering that the major group indicators present the average change across hundreds to thousands of species, it is likely that the variation across species will be high. In order to determine whether the patterns observed were a result of this random variation or were representing common responses across species we recreated the indicators presented in Figure 1 of the main text but randomised the species

within each major group. So, for each of the four major groups the same total number of species within it was maintained, but the species identified were randomly selected from the complete species pool. This species randomisation and indicator generation was carried out 12 times to see if the patterns of change observed could be recreated from the randomly selected set of species. Extended Data Figure 3 shows the analogue of Figure 1 in the main text, for each of the 12 alternative data sets obtained by randomisation in this way. For most of them, all four groups show very similar trends in occupancy that are similar to the overall increase of 11% reported in the main text: the fifth alternative dataset shows a separation into two pairs of groups, but none of these alternative datasets shows structures similar to those presented in Figure 1. This provides reassurance that those structures are indeed associated with genuine inter-group differences and cannot be attributed merely to random inter-species variation.

Posterior predictive checking

This section presents the results of some diagnostics that have been used to check our models' ability to reproduce selected features of the observations as aggregated over species groups: these checks have been carried out to provide some reassurance that the models are sufficiently flexible and realistic to capture the structures seen in the data, and hence to support the use of the models to make statements about long-term changes in occupancy.

Posterior predictive checking is a Bayesian technique that is designed to assess how well a model reproduces features of a data set[59]. The basic principle is, having fitted a model to a set of observations, y, to generate from the model a corresponding data vector, y_{rep}. If the data is appropriately parameterized then y and y_{rep} should be similar in some sense.

In the case of an occupancy-detection model, the observations for each species' model is a vector containing an entry 0 or 1 for each relevant visit in the database — 1 if the species was reported, 0 otherwise. 'Similarity' is defined by comparing relevant properties of the observed and simulated data vectors.

In our models, the observed detections y are assumed to be generated probabilistically and therefore the observed summary $T(y)$ is also drawn from some probability distribution: formally, it's the realised value of a random variable $T(Y)$, where Y is a random vector whose joint distribution is specified by the model. We don't know exactly what this distribution is because, even if the model structure is correct, we don't know the parameter values exactly. However, we do have a posterior distribution for the parameters. If we draw repeated samples from this posterior distribution and, for each sample, use the model to (a) generate a synthetic data set y_{rep} (b) calculate the corresponding summary $T(y_{rep})$, then we can build up a collection of samples from a distribution of $T(\cdot)$ that accounts both for the randomness in the model and for the parameter uncertainty. Such a distribution is called a "posterior predictive distribution". The observed summary $T(y)$ can then be compared with the posterior predictive distribution as a check on model performance.

In the present context, for each species the replications y_{rep} need to be generated to mimic as closely as possible the process that generated the data under the model: the same sites, same numbers of visits and associated list lengths. Two separate summary measures $T(y)$ were calculated for each group of species. The first was the overall proportion of sites with a detection, averaged over all species and years; and the second was the variance in the annual mean proportion of detections for the group (see below). This choice enables us to check the model's ability to reproduce features at the same level of group aggregation as the main analyses in the paper: moreover, the variance in annual mean detections is a measure of interannual variation which is related to the indices of change in the paper. Of course, it is not possible to compare the modelled occupancies with observations, because occupancy is not observed; since detection is conditional on occupancy however, we may have some confidence that models represent properties of the aggregated occupancies reasonably well if they can represent the corresponding properties of aggregated detections. We implemented the following protocol for each taxonomic group:

1. For each of V visits within each species' model, extract 99 samples from the posterior distribution of the probability that an observation was made on that visit. In practical terms, this probability is the product of the true (unknown) occupancy, z_{it}, and the detection probability, p_{itv} (see equation 4 in Outhwaite *et al* 2019).
2. Use each of the 99 sets of probabilities to sample a vector of potential observations under the model, by treating each visit as a potential Bernoulli trial. These vectors are 99 realisations of y_{rep}.
3. For each realisation y_{rep} and for each year (1970-2015), calculate the annual proportion of sites in which the species was recorded: denote this proportion, for species s and year t, by $T_{st}(y_{rep})$.
4. Calculate $T_{gt}(y_{rep})$ for each replicate dataset as the mean of $T_{st}(y_{rep})$ across species in each taxonomic group.
5. Calculate the mean across years as $T_{gm}(y_{rep})$ for each replicate dataset.
6. Calculate $T_{gv}(y_{rep})$ as the variance across years in $T_{gt}(y_{rep})$ for each replicate dataset.
7. Calculate the observed mean proportion of sites with records, $T_{gm}(y)$, and the variance across years $T_{gv}(y)$ for each replicate dataset.
8. Summarise the distribution of $T_{gm}(y)$ and $T_{gv}(y)$ as the mean and 95% credible intervals to demonstrate the variation in summary measures that can reasonably be expected under the model.

Computational limitations made this exercise unfeasible for four groups (Bryophyes, Dragonflies, Moths and Lichens). Molluscs are treated as a single group for this exercise since models were run as a complete group, but species were split into freshwater and terrestrial for the main analysis.

The mean proportion of sites with records is very well predicted on average, although rather uncertain in many taxonomic groups (Extended Data Figure 4).

The interannual variability is estimated very precisely by the model for most species (posterior predictive intervals are narrow, see Extended Data Figure 5), and shows excellent agreement with the observations for almost all taxonomic groups. There are a few taxa for which the observed value does not fall within the range of the posterior predictive interval, although in absolute terms the discrepancies are sufficiently small as not to compromise the main messages in the paper: the explanation for these discrepancies is either that the model is slightly biased for these species, or that the predictive uncertainties have been underestimated so that the intervals are slightly too narrow.

Author Contributions:

NJBI, BC and RDG conceived the study. CLO extracted and analysed the data and drafted the manuscript. REC determined the composite indicator method. All authors contributed to the writing and editing of the manuscript.

Data Availability:

The dataset analysed as a part of this study is publicly available from The Environmental Information Data Centre (EIDC) [30]. Additional information is supplied within the associated R package *UKBiodiversity* which is available from GitHub (https://github.com/CharlieOuthwaite/UKBiodiversity) and Data Descriptor [29].

Code Availability:

The code used to analyse the data is available from GitHub within an R package: *UKBiodiversity* (https://github.com/CharlieOuthwaite/UKBiodiversity).

Competing interests:

The authors declare no competing interests.

Acknowledgements:

We would like to acknowledge the input of the recording schemes and societies that contributed data to the dataset underlying this study and for their input in interpreting group level change. These schemes are: the Aquatic Heteroptera Recording Scheme, the Bees, Wasps and Ants Recording Society, the British Arachnological Society, Spider Recording Scheme, the British Bryological Society, the British Dragonfly Society: Dragonfly Recording Network, the British Lichen Society, the British Myriapod and Isopod Group: Centipede Recording Scheme, British Myriapod and Isopod Group: Millipede Recording Scheme, the Chrysomelidae Recording Scheme, the Conchological Society of Great Britain and Ireland, the Dipterists Forum: Cranefly Recording Scheme, the Dipterists Forum: Empididae, Hybotidae & Dolichopodidae Recording Scheme, the Dipterists Forum: Fungus Gnat Recording Scheme, the Dipterists Forum: Hoverfly Recording Scheme, the Gelechiid Recording Scheme, the

Grasshoppers and Related Insects Recording Scheme, the Ground Beetle Recording Scheme, the Lacewings and Allies Recording Scheme, the National Moth Recording Scheme, the Riverfly Recording Schemes: Ephemeroptera, the Riverfly Recording Schemes: Plecoptera, the Riverfly Recording Schemes: Trichoptera, the Soldier Beetles, Jewel Beetles and Glow-worms Recording Scheme, the Soldierflies and Allies Recording Scheme, the Staphylinidae Recording Scheme, the Terrestrial Heteroptera Recording Scheme - Plant bugs and allied species, the Terrestrial Heteroptera Recording Scheme - Shield bugs and allied species, the UK Ladybird Survey and the Weevil and Bark Beetle Recording Scheme. The authors would like to thank Georgina Mace, whose advice and comments on previous versions of this manuscript greatly improved the study. We are grateful to Tom August, Gary Powney, Jonathan Silvertown & Richard Pearson for advice and comments on draft manuscripts. Thank you also to Jamie Cranston for supplying a list of recent colonist species in the UK. This work was funded by the Natural Environment Research Council (NERC), award number NE/L008823/1 and was supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability. Finally, we would like to thank the three reviewers for their insightful comments on a previous draft of this manuscript.

References:

1. Gregory, R. D. *et al.* Developing indicators for European birds. *Philos. Trans. R. Soc.* **360**, 269–288 (2005).

2. McRae, L., Deinet, S. & Freeman, R. The diversity-weighted living planet index: Controlling for taxonomic bias in a global biodiversity indicator. *PLoS One* **12**, e0169156 (2017).

3. Brereton, T., Roy, D. B., Middlebrook, I., Botham, M. & Warren, M. The development of butterfly indicators in the United Kingdom and assessments in 2010. *J. Insect Conserv.* **15**, 139–151 (2010).

4. Van Swaay, C. A. M. *et al. The European Butterfly Indicator for Grassland species 1990-2013. Report VS2015.009*. (2015).

5. Tittensor, D. P. *et al.* A mid-term analysis of progress toward international biodiversity targets. *Science (80-.).* **346**, 241–244 (2014).

6. Westgate, M. J., Barton, P. S., Lane, P. W. & Lindenmayer, D. B. Global meta-analysis reveals low consistency of biodiversity congruence relationships. *Nat. Commun.* **5**, 3899 (2014).

7. Rodrigues, A. S. L. & Brooks, T. M. Shortcuts for Biodiversity Conservation Planning: The Effectiveness of Surrogates. *Annu. Rev. Ecol. Evol. Syst.* **38**, 713–737 (2007).

8. Hambler, C. & Speight, M. R. Extinction rates and butterflies. *Science* **305**, 1563–5; author reply 1563-5 (2004).

9. van Strien, A. J. *et al.* Modest recovery of biodiversity in a western European country: The

480 Living Planet Index for the Netherlands. *Biol. Conserv.* **200**, 44–50 (2016).

481 10. Hallmann, C. A. *et al.* More than 75 percent decline over 27 years in total flying insect
482 biomass in protected areas. *PLoS One* **12**, e0185809 (2017).

483 11. Sánchez-Bayo, F. & Wyckhuys, K. A. G. Worldwide decline of the entomofauna: A review of
484 its drivers. *Biol. Conserv.* **232**, 8–27 (2019).

485 12. Powney, G. D. *et al.* Widespread losses of pollinating insects in Britain. *Nat. Commun.* **10**,
486 1018 (2019).

487 13. Simmons, B. I. *et al.* Worldwide insect declines: An important message, but interpret with
488 caution. *Ecol. Evol.* **9**, 3678–3680 (2019).

489 14. Thomas, C. D., Jones, T. H. & Hartley, S. E. "Insectageddon": A call for more robust data and
490 rigorous analyses. *Glob. Chang. Biol.* (2019). doi:10.1111/gcb.14608

491 15. Burns, F. *et al.* Agricultural Management and Climatic Change Are the Major Drivers of
492 Biodiversity Change in the UK. *PLoS One* **11**, e0151595 (2016).

493 16. Burns, F. *et al.* An assessment of the state of nature in the United Kingdom: A review of
494 findings, methods and impact. *Ecol. Indic.* **94**, 226–236 (2018).

495 17. Hayhow, D. *et al. The state of the UK's birds 2017*. (2017).

496 18. Fox, R. *et al. The State of the UK's Butterflies 2015*. (2015).

497 19. Maskell, L. C., Smart, S. M., Bullock, J. M., Thompson, K. & Stevens, C. J. Nitrogen
498 deposition causes widespread loss of species richness in British habitats. *Glob. Chang. Biol.*
499 **16**, 671–679 (2010).

500 20. Barlow, K. E. *et al.* Citizen science reveals trends in bat populations: The National Bat
501 Monitoring Programme in Great Britain. *Biol. Conserv.* **182**, 14–26 (2015).

502 21. Pocock, M. J. O., Roy, H. E., Preston, C. D. & Roy, D. B. The Biological Records Centre: A
503 pioneer of citizen science. *Biol. J. Linn. Soc.* **115**, 475–493 (2015).

504 22. Isaac, N. J. B. & Pocock, M. J. O. Bias and information in biological records. *Biol. J. Linn.*
505 *Soc.* **115**, (2015).

506 23. Boakes, E. H. *et al.* Distorted views of biodiversity: Spatial and temporal bias in species
507 occurrence data. *PLoS Biol.* **8**, (2010).

508 24. Maldonado, C. *et al.* Estimating species diversity and distribution in the era of Big Data: to
509 what extent can we trust public databases? *Glob. Ecol. Biogeogr.* **24**, 973–984 (2015).

510 25. Outhwaite, C. L. *et al.* Prior specification in Bayesian occupancy modelling improves analysis
511 of species occurrence data. *Ecol. Indic.* **93**, 333–343 (2018).

512 26. Isaac, N. J. B., van Strien, A. J., August, T. a, de Zeeuw, M. P. & Roy, D. B. Statistics for
513 citizen science: extracting signals of change from noisy ecological data. *Methods Ecol. Evol.* **5**,
514 1052–1060 (2014).

515 27. van Strien, A. J., van Swaay, C. A. M. & Termaat, T. Opportunistic citizen science data of
516 animal species produce reliable estimates of distribution trends if analysed with occupancy
517 models. *J. Appl. Ecol.* **50**, 1450–1458 (2013).

518 28. Termaat, T. *et al.* Distribution trends of European dragonflies under climate change. *Divers.*
519 *Distrib.* (2019). doi:10.1111/ddi.12913

520 29. Outhwaite, C. L. *et al.* Annual estimates of occupancy for bryophytes, lichens and
521 invertebrates in the UK (1970-2015). (2019). doi:10.5285/0ec7e549-57d4-4e2d-b2d3-

2199e1578d84

30. Dornelas, M. *et al.* A balance of winners and losers in the Anthropocene. *Ecol. Lett.* ele.13242 (2019). doi:10.1111/ele.13242

31. Gregory, R. & van Strien, A. Wild bird indicators: using composite population trends of birds as measures of environmental health. *Ornithol. Sci.* **9**, 3–22 (2010).

32. Newbold, T. *et al.* Global effects of land use on local terrestrial biodiversity. *Nature* **520**, 45–50 (2015).

33. Pescott, O. L. *et al.* Air pollution and its effects on lichens, bryophytes, and lichen-feeding Lepidoptera: Review and evidence from biological records. *Biol. J. Linn. Soc.* **115**, 611–635 (2015).

34. Saal, D. S. & Parker, D. The impact of privatization and regulation on the water and sewerage industry in England and Wales: a translog cost function model. *Manag. Decis. Econ.* **21**, 253–268 (2000).

35. Vaughan, I. P. & Ormerod, S. J. Large-scale, long-term trends in British river macroinvertebrates. *Glob. Chang. Biol.* **18**, 2184–2194 (2012).

36. Vaughan, I. P. & Gotelli, N. J. Water quality improvements offset the climatic debt for stream macroinvertebrates over twenty years. *Nat. Commun.* **10**, 1956 (2019).

37. Langford, T. E. L., Shaw, P. J., Ferguson, A. J. D. & Howard, S. R. Long-term recovery of macroinvertebrate biota in grossly polluted streams: Re-colonisation as a constraint to ecological quality. *Ecol. Indic.* **9**, 1064–1077 (2009).

38. Balmford, A. & Knowlton, N. Why Earth Optimism? *Science* **356**, 225 (2017).

39. Antrop, M. Why landscapes of the past are important for the future. *Landsc. Urban Plan.* **70**, 21–34 (2005).

40. Robinson, R. A. & Sutherland, W. J. Post-war changes in arable farming and biodiversity in Great Britain. *J. Appl. Ecol.* **39**, 157–176 (2002).

41. Mihoub, J. B. *et al.* Setting temporal baselines for biodiversity: The limits of available monitoring data for capturing the full impact of anthropogenic pressures. *Sci. Rep.* **7**, 41591 (2017).

42. Soga, M. & Gaston, K. J. Shifting baseline syndrome: causes, consequences, and implications. *Front. Ecol. Environ.* **16**, 222–230 (2018).

43. Thomas, C. D. *Inheritors of the Earth: how nature is thriving in an age of extinction*. (Hachette UK, 2017).

44. Dornelas, M. *et al.* Assemblage time series reveal biodiversity change but not systematic loss. *Science* **344**, 296–9 (2014).

45. Lister, B. C. & Garcia, A. Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proc. Natl. Acad. Sci.* 201722477 (2018). doi:10.1073/pnas.1722477115

46. Mace, G. M., Collen, B., Fuller, R. A. & Boakes, E. H. Population and geographic range dynamics: implications for conservation planning. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 3743–3751 (2010).

47. Gaston, K. J. & Fuller, R. A. The sizes of species' geographic ranges. *J. Appl. Ecol.* **46**, 1–9 (2009).

48. Seibold, S. *et al.* Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature* **574**, 671–674 (2019).

49. Bart, J. & Klosiewski, S. P. Use of Presence-Absence to Measure Changes in Avian Density. *J. Wildl. Manage.* **53**, 847 (1989).

50. Webb, T. J., Freckleton, R. P. & Gaston, K. J. Characterizing abundance-occupancy relationships: there is no artefact. *Glob. Ecol. Biogeogr.* **21**, 952–957 (2012).

51. Buckley, H. L. & Freckleton, R. P. Understanding the role of species dynamics in abundance-occupancy relationships. *J. Ecol.* **98**, 645–658 (2010).

52. van Strien, A. J., van Swaay, C. A. M., van Strien-van Liempt, W. T. F. H., Poot, M. J. M. & WallisDeVries, M. F. Over a century of data reveal more than 80% decline in butterflies in the Netherlands. *Biol. Conserv.* **234**, 116–122 (2019).

53. Jetz, W. *et al.* Essential biodiversity variables for mapping and monitoring species populations. *Nat. Ecol. Evol.* (2019). doi:10.1038/s41559-019-0826-1

54. Schmeller, D. S. *et al.* An operational definition of essential biodiversity variables. *Biodivers. Conserv.* (2017). doi:10.1007/s10531-017-1386-9

55. McGill, B. J., Dornelas, M., Gotelli, N. J. & Magurran, A. E. Fifteen forms of biodiversity trend in the Anthropocene. *Trends Ecol. Evol.* (2014). doi:10.1016/j.tree.2014.11.006

56. Buckland, S. T., Magurran, A. E., Green, R. E. & Fewster, R. M. Monitoring change in biodiversity through composite indices. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 243–54 (2005).

57. Faraway, J. J. *Linear models with R.* (Chapman & Hall / CRC, 2009).

58. Outhwaite, C. L. *et al.* Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970-2015. *Sci. data* **6**, 259 (2019).

59. Gelman, A. *et al. Bayesian Data Analysis*. (CRC Press, 2014).

Figure Legends:

Figure 1: Composite estimates of average annual occupancy of four groups of species. *Values are scaled to 100 in 1970. Coloured lines show the average response as the geometric mean occupancy and the shaded area represents the 95% credible intervals of the posterior distribution of the geometric mean.* n *denotes the number of species contributing to each group. Uncertainty for each year is expressed relative to the 1970 baseline. Change metrics reported in the text account for uncertainty in both the first and last year of the series.*

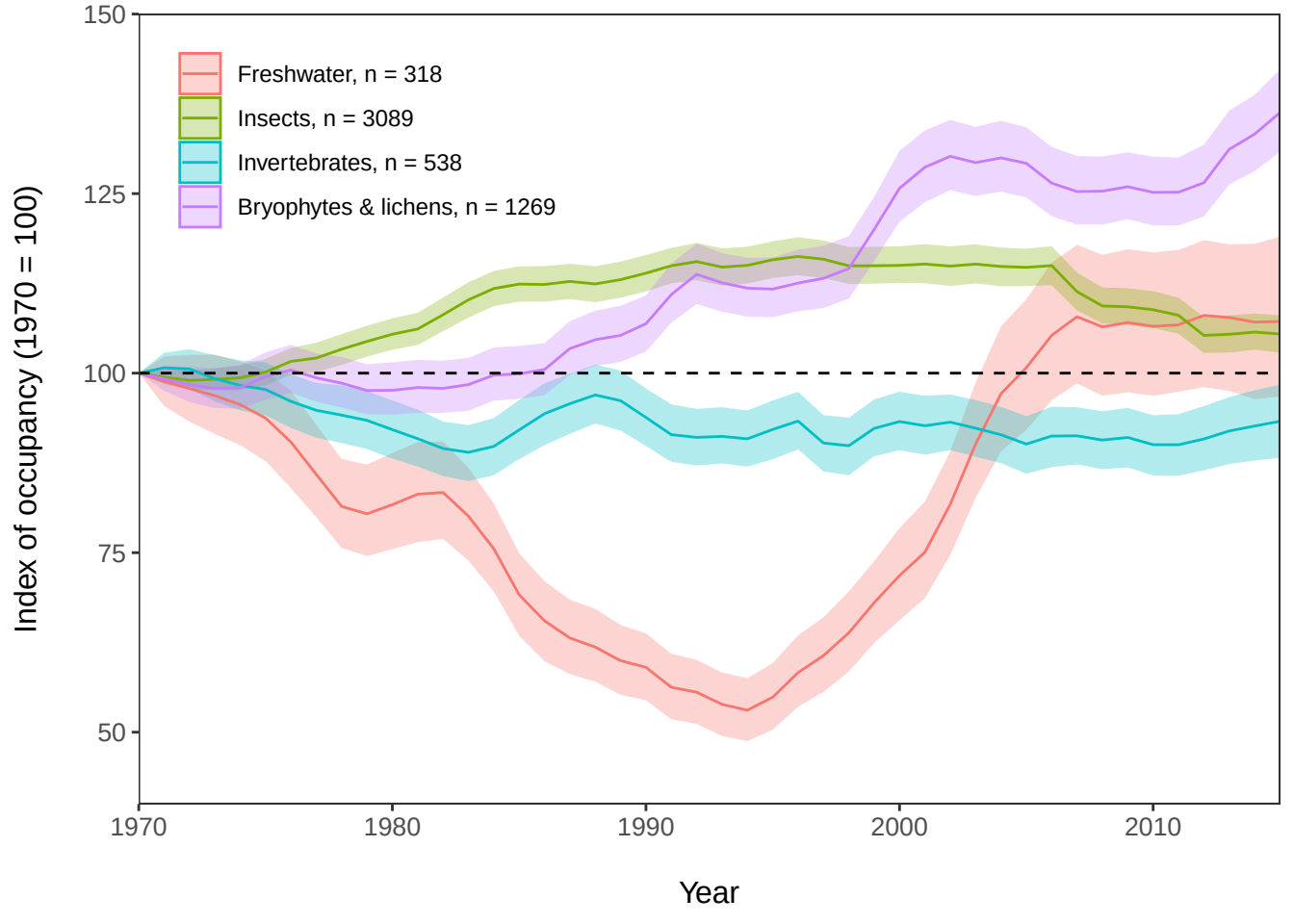
Figure 2: Absolute change in geometric mean occupancy during the first (1970-1992) and second (1993–2015) halves of the time series for each major group. *Each boxplot represents the posterior distribution of overall absolute change in occupancy within the group, over the relevant time period. The centre of the boxplot represents the median of the distribution with lower and upper hinges corresponding to the 25th and 75th percentiles. The whiskers represent the 95% credible intervals.*

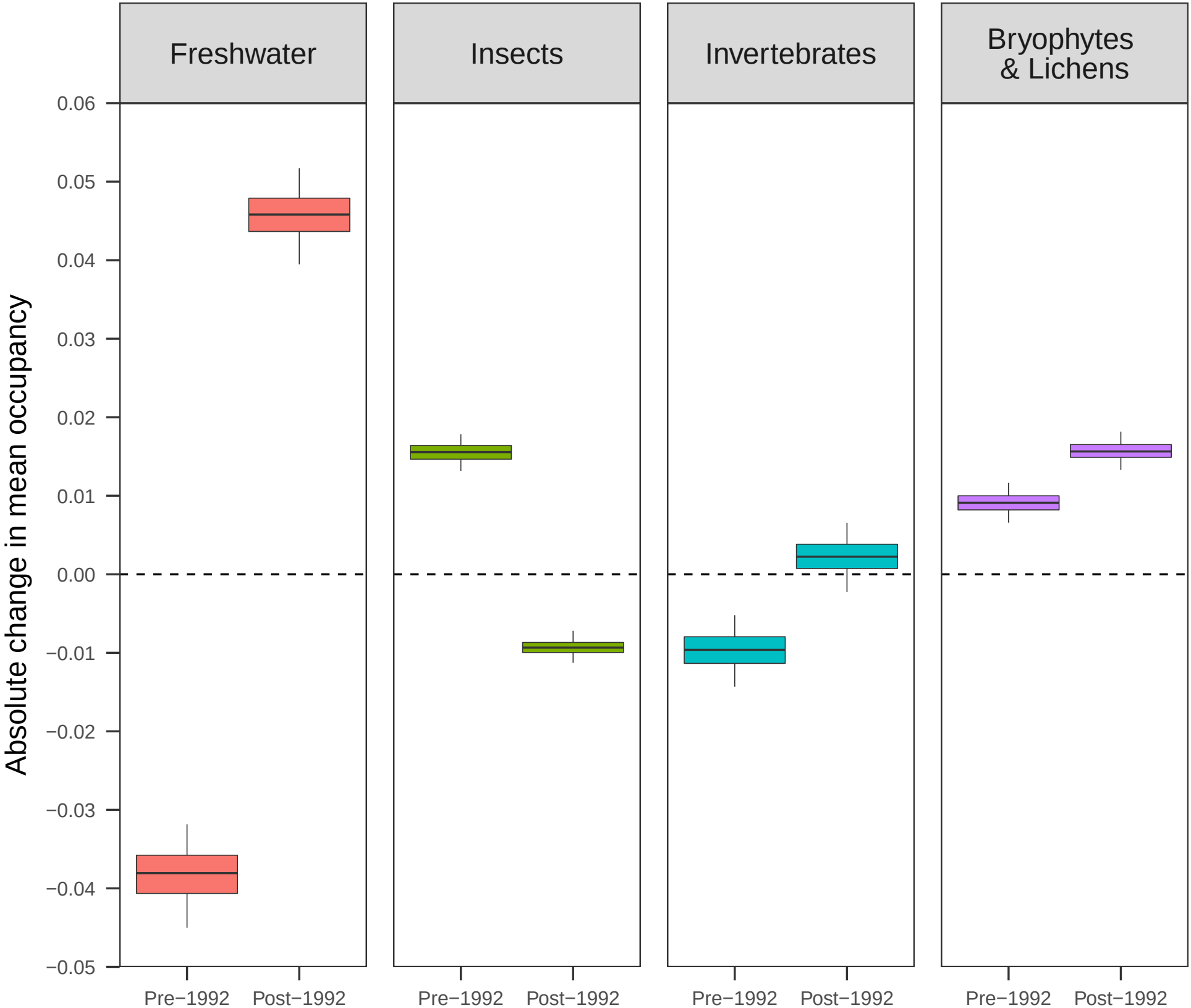
Figure 3: Composite estimates of two quantiles of annual occupancy across the four major groups. *Two quantiles were chosen to represent varying levels of occupancy: common or widespread (0.75 -*

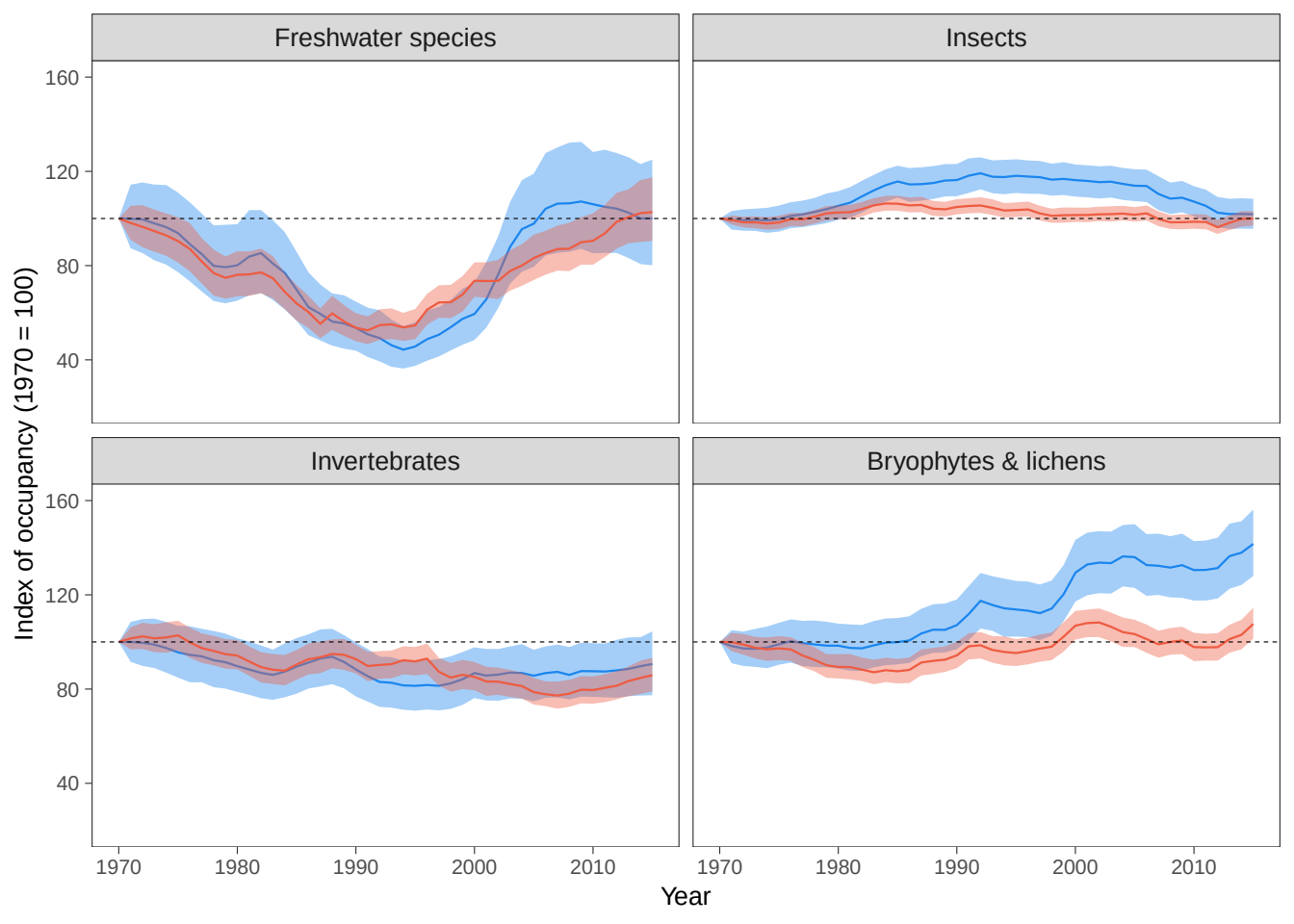
603 *green) and rare or localised occupancy (0.25 - purple). These quantiles of occupancy were assessed*
604 *each year to show how rareness and commonness changed over time for each group. In each case, the*
605 *shaded area delimits the 95% credible intervals.*

606 ***Figure 4: Composite estimates of average annual occupancy of each taxonomic subgroup***. *Taxa*
607 *within the freshwater and insect groups are displayed across multiple panels to aid visibility. Values*
608 *are scaled to 100 in 1970. Coloured lines show the average response as the geometric mean occupancy*
609 *and the shaded area represents the 95% credible intervals of the posterior distribution of annual*
610 *occupancy estimates. The width of the credible intervals is a function of the number of species within*
611 *the group. Note that y-axes vary.*

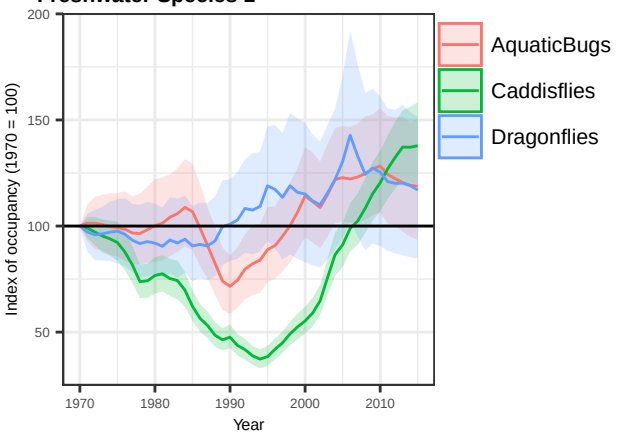
612 ***Figure 5: Heat map of comparison between each species' average occupancy estimate across the***
613 ***entire period and its' average annual growth rate (expressed as a percentage of the initial occupancy,***
614 ***see Methods) for each of the four major groups.*** *Brighter shades represent a greater number of species*
615 *within that hexagon. The greater the average occupancy value on the y- axis, the more common the*
616 *species. Hexagons to the left of the vertical, dashed line (growth rate = 0) include species with a*
617 *negative annual growth rate, those on the right have a positive annual growth rate. Six extreme positive*
618 *growth rates are not shown.*


Freshwater, n = 318
Insects, n = 3089
Invertebrates, n = 538
Bryophytes & lichens, n = 1269
Index of occupancy (1970 = 100)
150
125
100
75
50
1970
1980
1990
2000
2010
Year

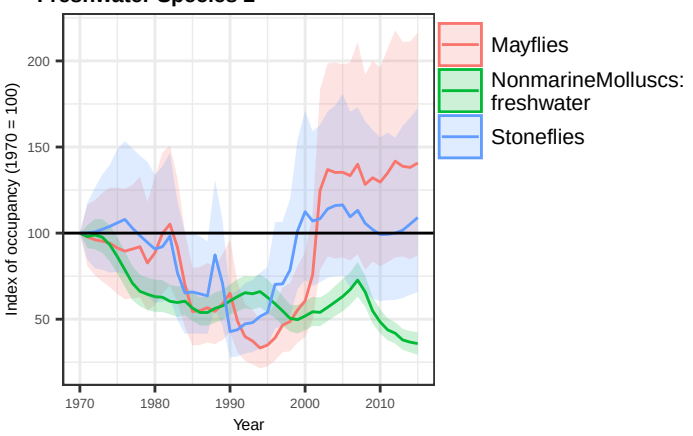

Freshwater
Insects
Invertebrates
Bryophytes & Lichens
Absolute change in mean occupancy
0.06
0.05
0.04
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
−0.04
−0.05
Pre−1992
Post−1992
Pre−1992
Post−1992
Pre−1992
Post−1992
Pre−1992
Post−1992


Freshwater species
Insects
Invertebrates
Bryophytes & lichens
Index of occupancy (1970 = 100)
160
120
80
40
1970
1980
1990
2000
2010
Year

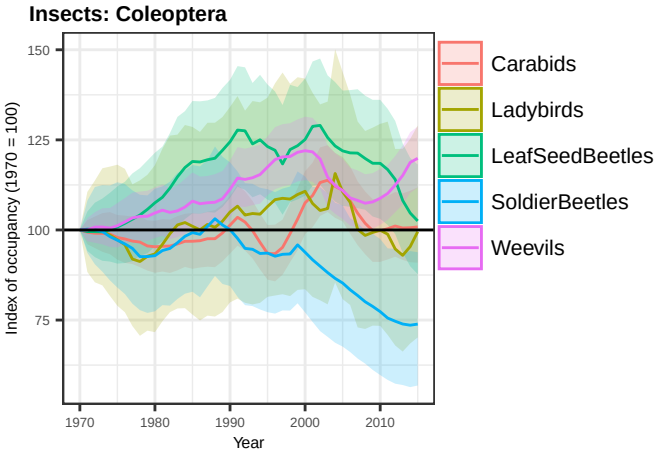
Freshwater Species 1



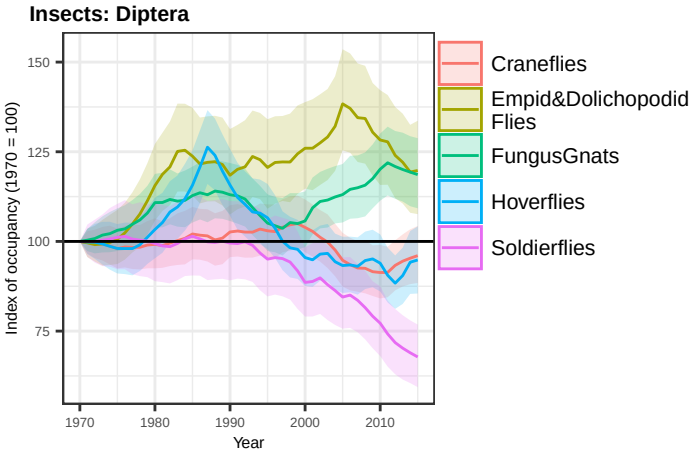
Freshwater Species 2



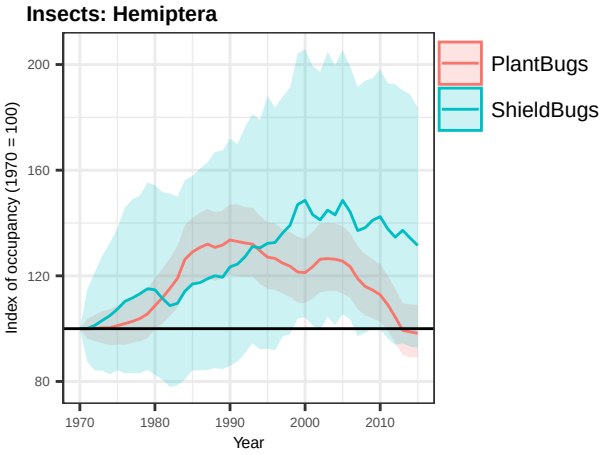
Insects: Coleoptera



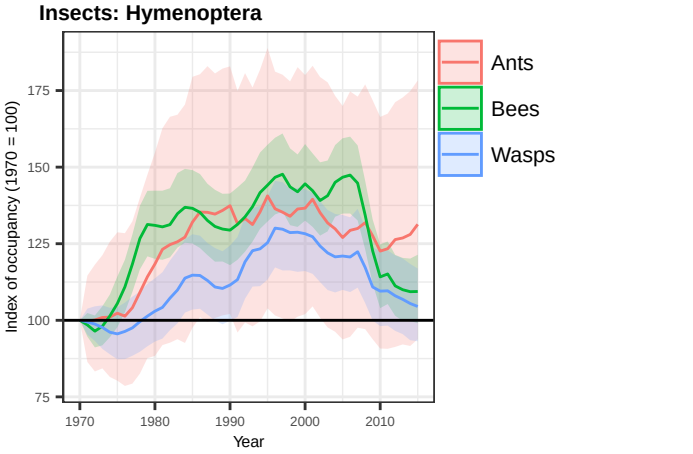
Insects: Diptera



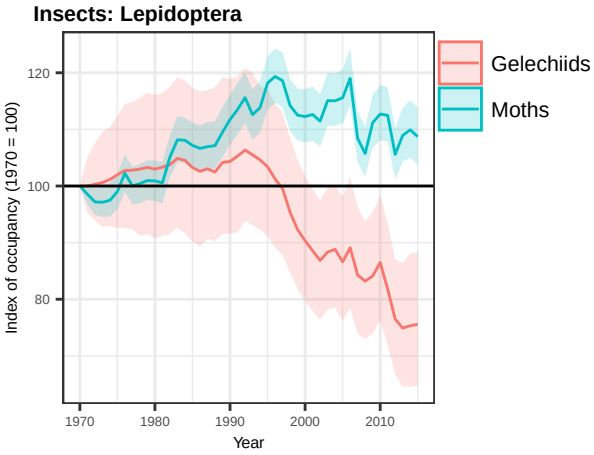
Insects: Hemiptera



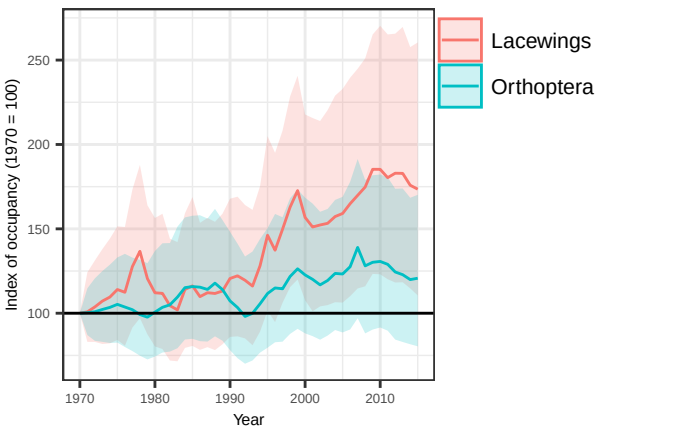
Insects: Hymenoptera



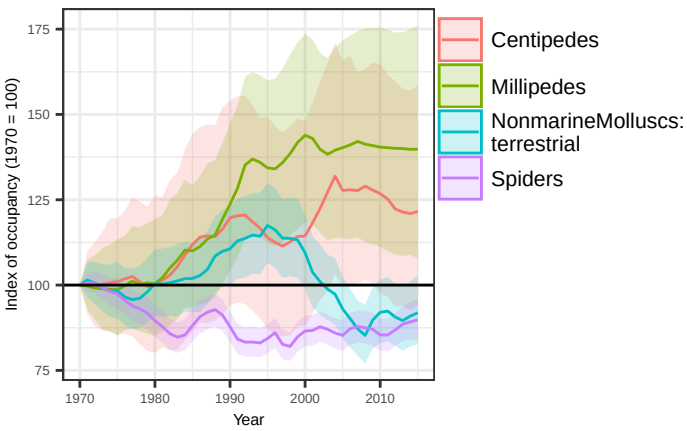
Insects: Lepidoptera



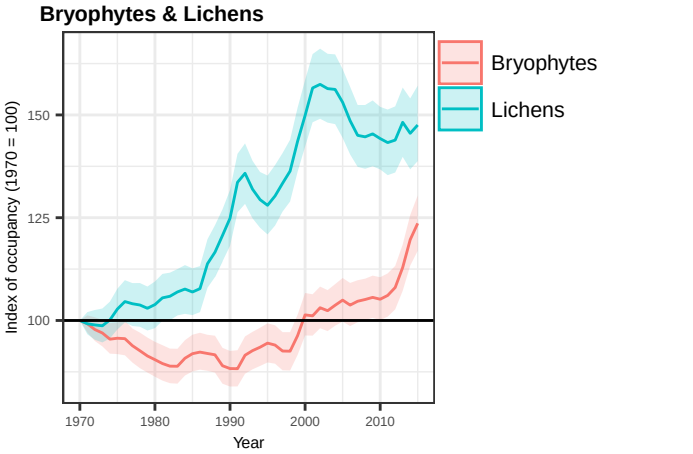
Insects: other

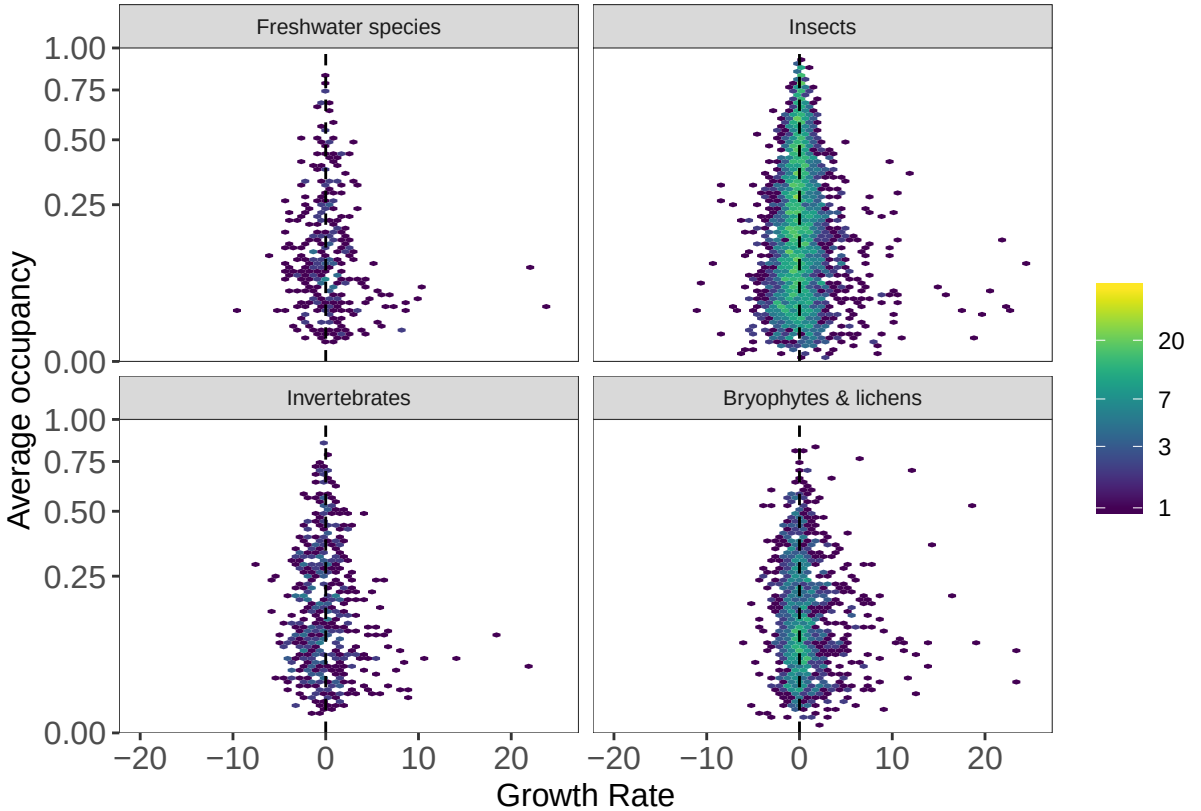


Invertebrates

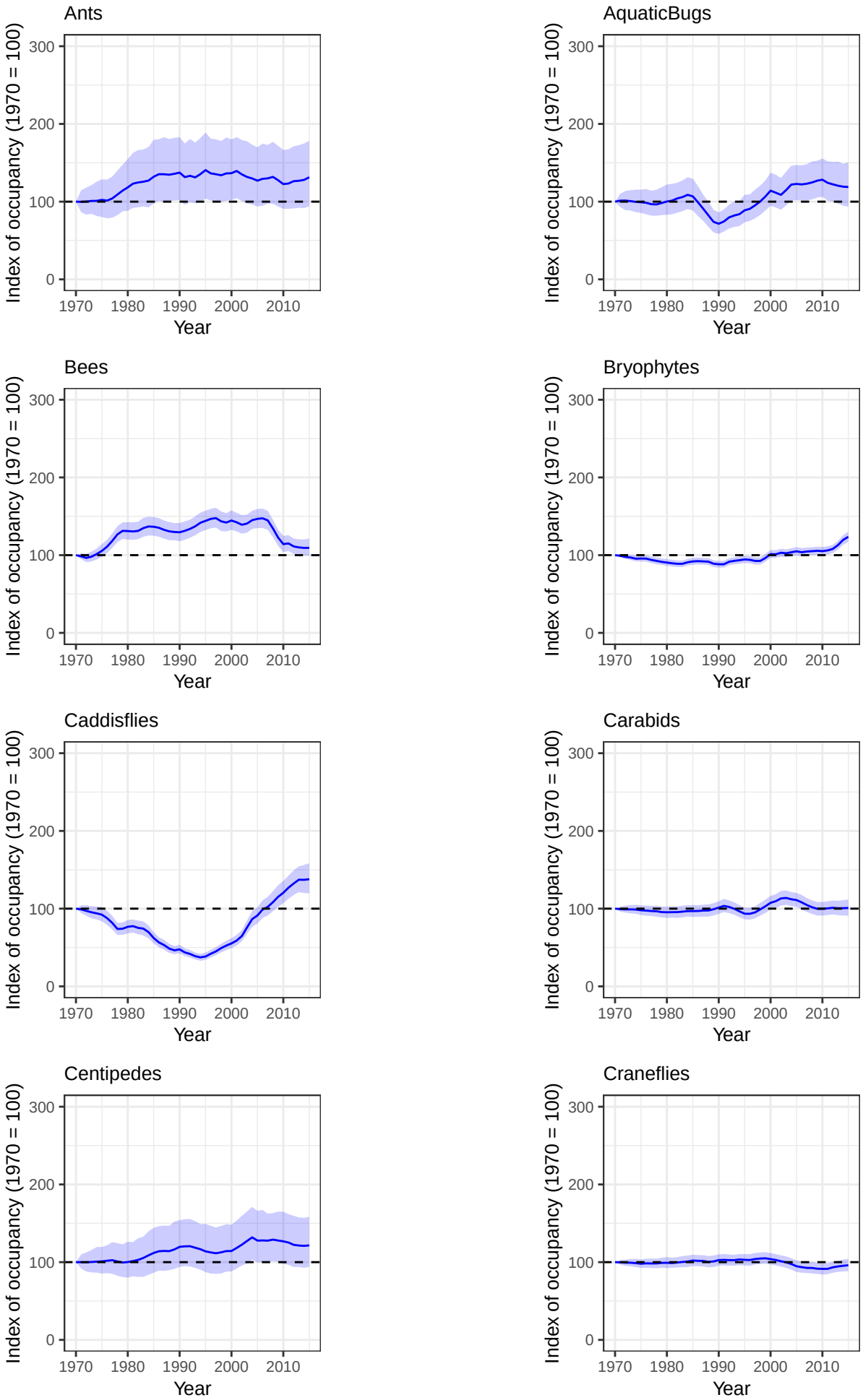


Bryophytes & Lichens

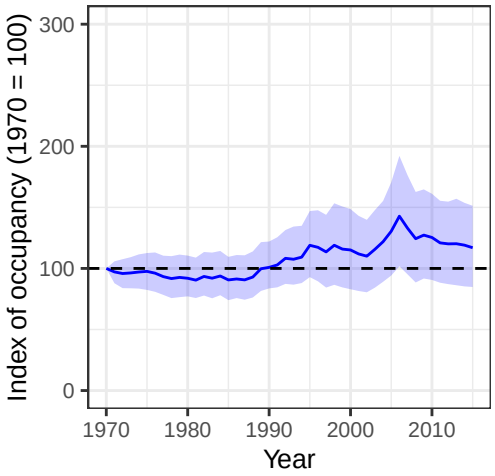



Freshwater species
Insects
Invertebrates
Bryophytes & lichens
Average occupancy
1.00
0.75
0.50
0.25
0.00
−20
−10
0
10
20
Growth Rate
20
7
3
1

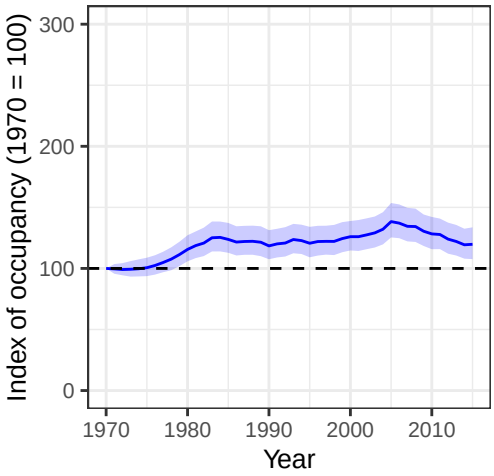
Supplementary Figure 1:
Average occupancy over time Values are scaled to 100 in 1970.
Coloured lines show the average response as the geometric mean occupancy
and the shaded area represents the 95% credible intervals
of the posterior distribution of the geometric mean



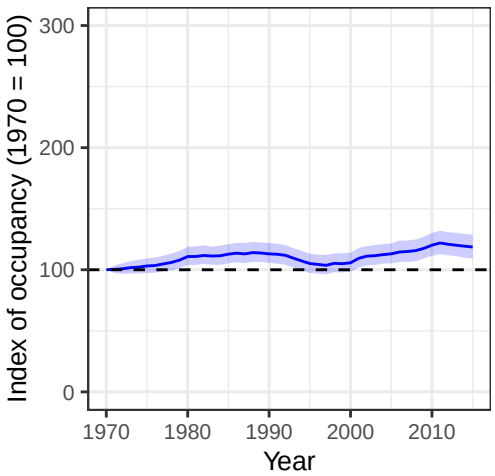
Dragonflies



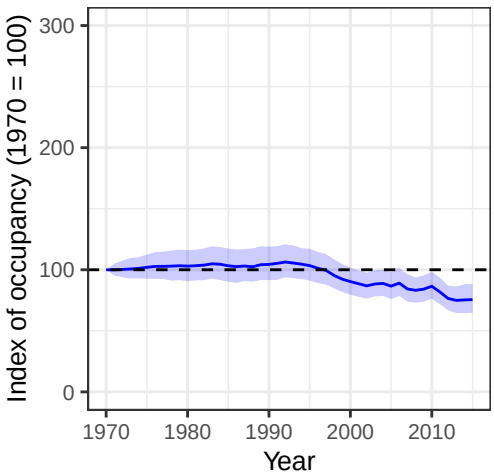
Empid&DolichopodidFlies



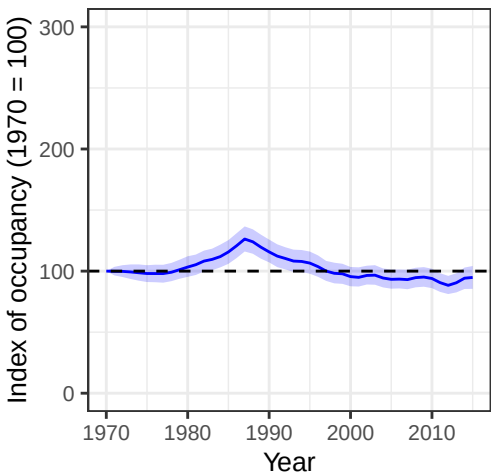
FungusGnats



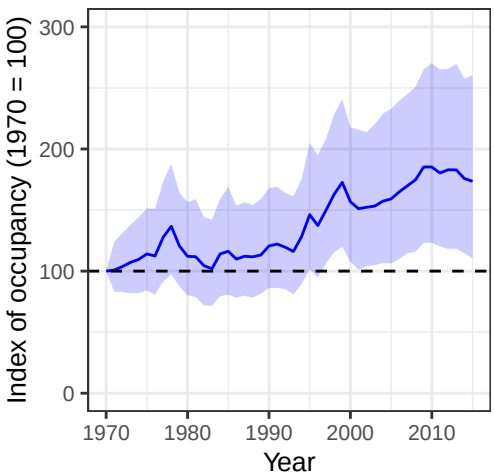
Gelechiids



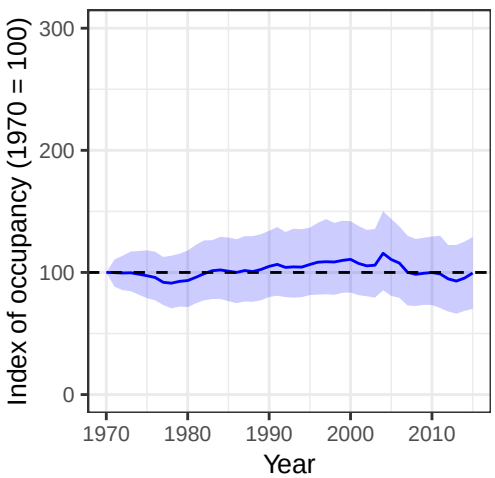
Hoverflies



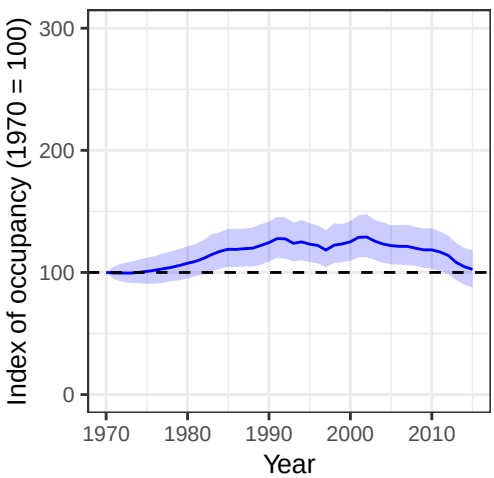
Lacewings



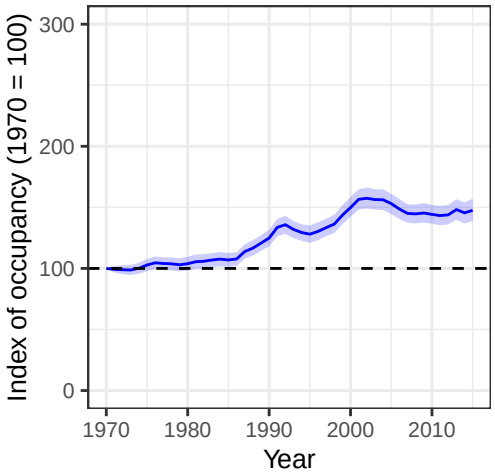
Ladybirds



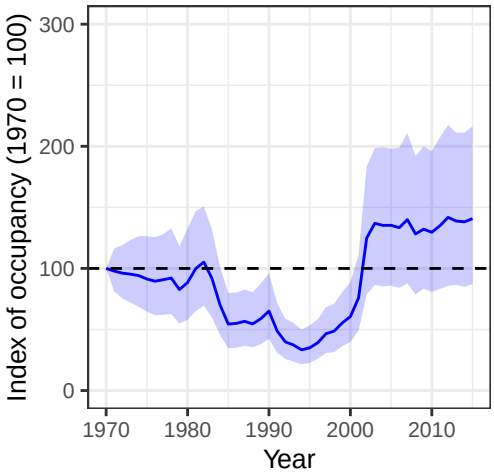
LeafSeedBeetles



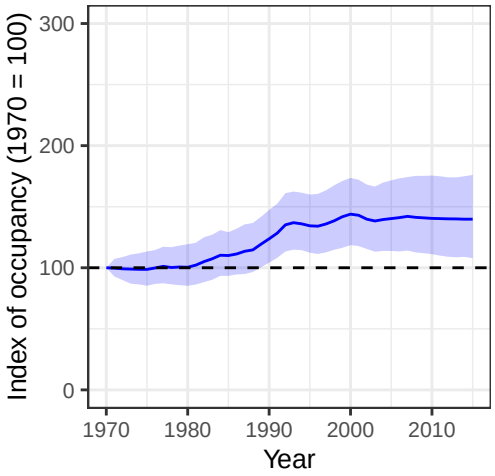
Lichens



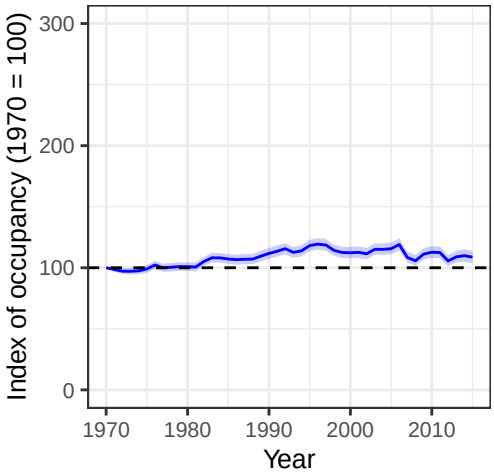
Mayflies



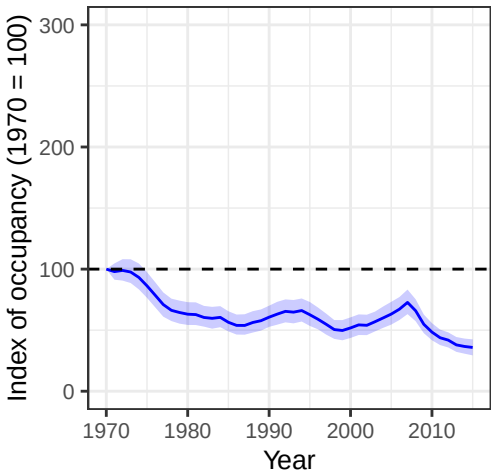
Millipedes



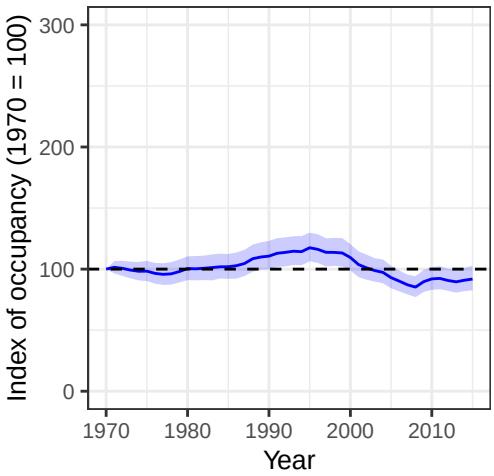
Moths



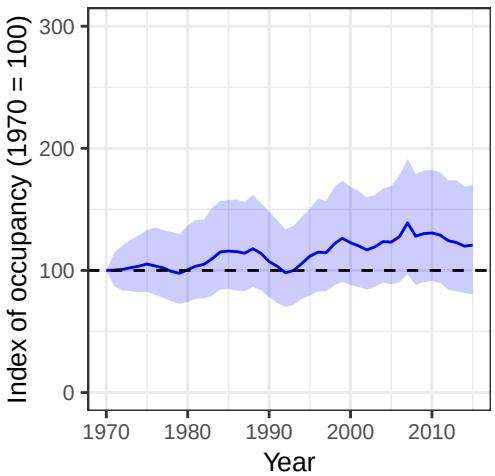
NonmarineMolluscs_freshwater



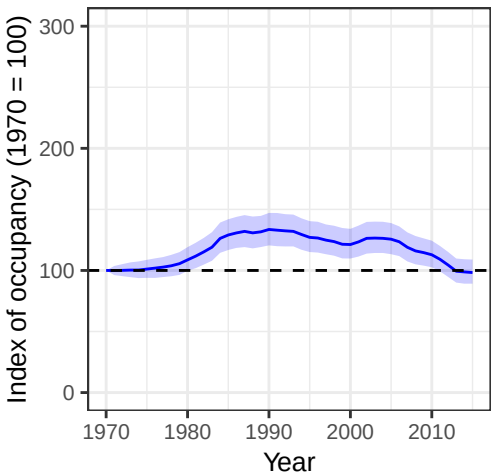
NonmarineMolluscs_terrestrial



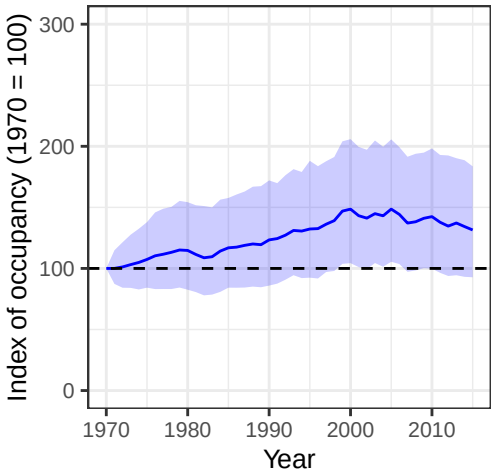
Orthoptera



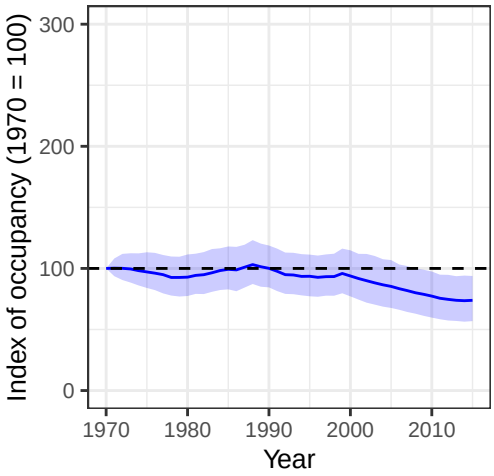
PlantBugs



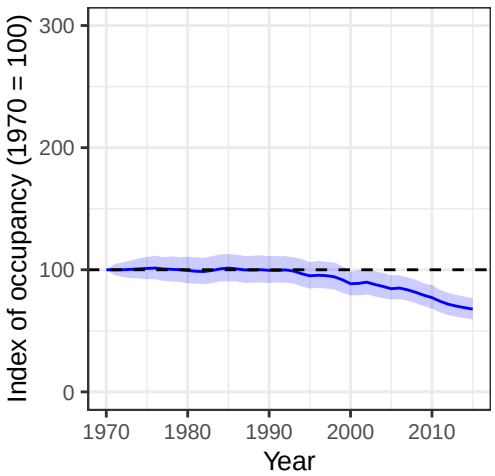
ShieldBugs



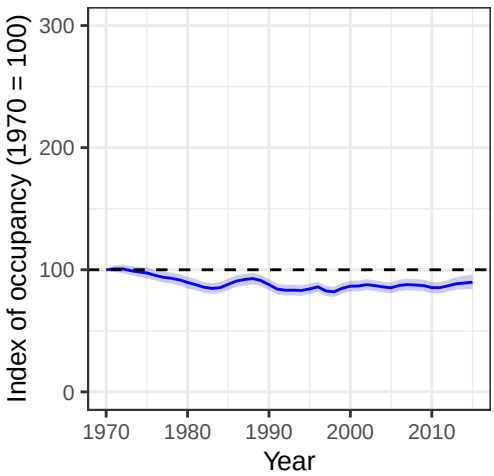
SoldierBeetles



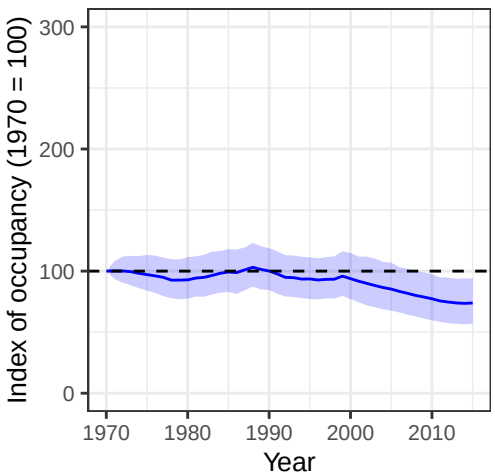
Soldierflies



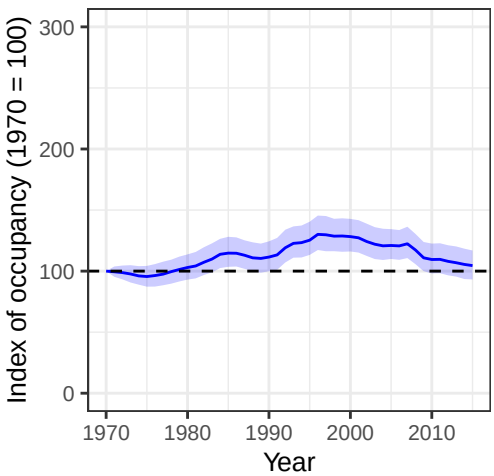
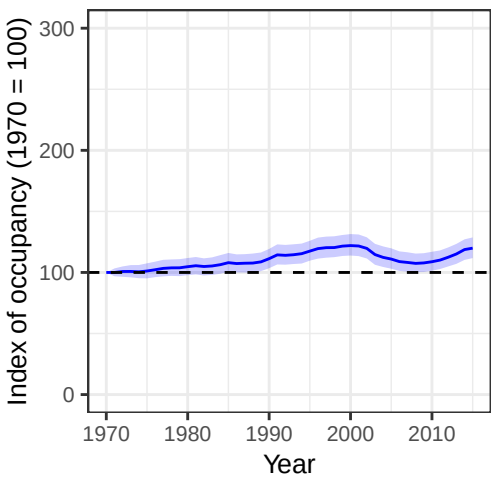
Spiders



SoldierBeetles



Wasps



Weevils