Ecography 34: 392–401, 2011
doi: 10.1111/j.1600-0587.2010.06554.x

Subject Editor: Miguel Araujo. Accepted 24 August 2010

Biodiversity change is scale-dependent: an example from Dutch and UK hoverflies (Diptera, Syrphidae)

Petr Keil, Jacobus C. Biesmeijer, Aat Barendregt, Menno Reemer and William E. Kunin

P. Keil (pkeil@seznam.cz), Dept of Ecology, Faculty of Science, Charles Univ., Vinicna 7, CZ-12844 Praha 2, Czech Republic. – J. C. Biesmeijer and W. E. Kunin, Inst. of Integrative and Comparative Biology, Faculty of Biological Sciences, Univ. of Leeds, Leeds LS2 9JT, UK. – A. Barendregt, Dept of Environmental Sciences – Copernicus Inst. Utrecht Univ., PO Box 80115, NL-3508 TC Utrecht, the Netherlands. – M. Reemer, European Invertebrate Survey – the Netherlands, Netherlands Centre for Biodiversity, Naturalis, Postbus 9517, NL-2300 RA Leiden, the Netherlands.

We test whether temporal change in species richness (ΔS [%]) is scale-dependent, using data on hoverflies from the UK and the Netherlands. We analysed ΔS between pre-1980 and post-1980 periods using 5 grid resolutions (10×10, 20×20, 40×40, 80×80 and 160×160 km). We also tested the effect of data quality and of unequal survey periods on ΔS estimates, and checked for spatial autocorrelation of ΔS estimates. Using data from equal survey periods, we found significant increases in hoverfly species richness in the Netherlands at fine scales, but no significant change at coarser scales indicating a decrease in beta diversity. In the UK, ΔS was negative at fine scale, near zero at intermediate scales, and positive at coarse scales, indicating that the degree of spatial beta diversity increased between the time periods. The use of unequal survey periods (using longer periods in the past to compensate for lower survey intensity) tended to inflate past species richness, biasing ΔS estimates downwards. High data quality thresholds sometimes obscured dynamics by reducing sample size, but never reversed trends. There was little spatial autocorrelation of ΔS, implying that local drivers (land use change or environmental noise) are important in dynamics of hoverfly diversity. A second, sample agglomeration approach to measure scaling resulted in greater noise in ΔS, obscuring the NL pattern, while still showing strong evidence of fine-scale richness loss in the UK. Our results indicate that explicit considerations of spatial (and temporal) scale are essential in studies documenting past biodiversity change, or projecting change into the future.

Ecological patterns are often dependent on the spatial scale of observation. This scale-dependence has recently attracted considerable attention (Storch et al. 2007). Recent studies have demonstrated that species distributions (Kunin 1998), density dependence (Jarosik and Lapchin 2001, Gunton and Kunin 2007, 2009), extinction risk (Hartley and Kunin 2003), ratios of native/exotic species (Davies et al. 2005) or migration and colonization rates (Menendez and Thomas 2000, Englund and Hamback 2007) manifest themselves differently across spatial scales, and it has long been appreciated that species richness is intrinsically scale-dependent (Arrhenius 1921, Rahbek and Graves 2001, Whittaker et al. 2001, Rahbek 2005). Such scale-specificity may be attributed to the scaling properties of environmental factors, biotic processes, and their interaction, which together result in differences in relative importance of various factors at different scales (Shmida and Wilson 1985, Wiens 1989). Species richness, for example, may be primarily driven by habitat structure and biotic interactions at a local scale, but climatic influences are thought to be more important at regional or continental scales (Willis and Whittaker 2002). Moreover, some factors, such as productivity, may show effects on biodiversity that differ in strength or even direction at different scales (Chase and Leibold 2002).

To date, most studies looking at patterns of biodiversity at multiple spatial scales have focused on static diversity distributions, i.e. examining a single snapshot of species richness pattern at a specific point in time (but see Hartley and Kunin 2003). However, there is increasing evidence of substantial temporal changes in species diversity and distributions – mostly alarming declines (Warren et al. 2001, Thomas et al. 2004, Biesmeijer et al. 2006, Hickling et al. 2006), but also increases due to climate impacts and human-assisted dispersal (e.g. of alien species; Preston et al. 2002, Stohlgren et al. 2003). Such reports typically assess only one spatial scale of observation.

While there is little direct evidence that changes in diversity are scale-dependent, there are reasons to suspect that change will be scale-dependent. The main drivers of biodiversity change are all arguably scale-specific in their likely impacts. Habitat fragmentation and subsequent metapopulation dynamics, for example, tend to result in the disproportionate loss of small, isolated populations

(Hanski 1999). Loss of such isolates involves substantial area loss at coarse scales, but relatively little loss at fine scales (as the patches lost are small). Climate change results in both range expansions and contractions, but evidence to date suggests that these two effects display strikingly different distributional scaling (Wilson et al. 2004, Pocock et al. 2006); ranges tend to expand in a spatially cohesive fashion (involving a substantial fine-scale gain as coarse-scale areas are colonised), but leave behind scattered populations as they contract (causing little loss at coarse scale despite substantial fine-scale losses). The impact of invasive alien species is also predicted to be scale-dependent (Rosenzweig 2001, Stohlgren et al. 2006); invasions may reduce very fine scale diversity (in some cases) and global scale diversity (if they lead to the extinction of native populations), but they almost invariably increase total diversity at intermediate spatial scales.

The scale-dependency of biodiversity change between two time periods can be described as a change in the slope (one metric of β diversity) of the species–area relationship (SAR) over time (Fig. 1). However, any attempt to formulate an a priori prediction about the precise scale-dependency of diversity changes would need either complex models of range dynamics or scenario-based projections (Thuiller et al. 2005, Jetz et al. 2007), in which the relative importance of different drivers of change was known. This is beyond the scope of our study as we restrict our goal to merely ask whether and how are changes in biodiversity scale-dependent.

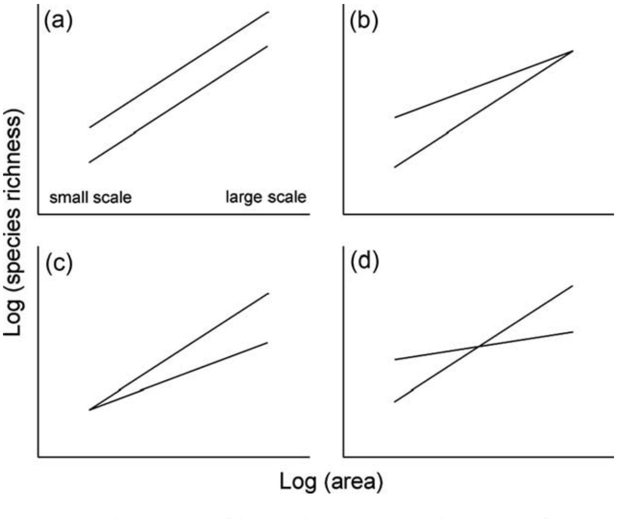


Figure 1. Illustration of how relative temporal change of species richness (ΔS) in time can depend on scale. Two lines in each plot represent two SARs measured in subsequent periods of time. Any scale-dependency of ΔS is caused only by change in the slope of SAR (which is β diversity). (a) ΔS is independent on spatial scale. (b–c) ΔS equals zero at one scale but is different from zero (either positive or negative) at the other scale. (d) ΔS equals zero at intermediate scale and differs from zero at large and small scales (ΔS have opposite directions at large and small scales). Note that we show the linear log-log SARs only for illustration. Scale-dependent ΔS will emerge not only when the two SARs differ in slope, but also when they differ in shape (e.g. when one is linear and the other non-linear in log-log). Such situation will always be interpretable as a change in β diversity, although not necessarily at all spatial scales. See also Supplementary material Fig. S1 for rough estimation of SARs for the UK and NL data.

The primary goal of this paper is to test the hypothesis that spatial scale influences the observed patterns of biodiversity change. We use two extensive data sets on hoverflies (Insecta: Diptera: Syrphidae) and assess the change during the second half of the 20th century in the NL and the UK at five grain resolutions (scales): 10 × 10, 20 × 20, 40 × 40, 80 × 80 and 160 × 160 km. Our secondary goal was to estimate the level of spatial autocorrelation of biodiversity change. Strong autocorrelation would indicate presence of a common factor influencing biodiversity at large scales (e.g. climate; Diniz-Filho et al. 2003), whereas weak (or lack of) autocorrelation would point to the importance of local factors (e.g. land use change) or to the presence of strong fine-scale environmental or demographic noise. Additionally, we assessed the potentially confounding effect of data quality on biodiversity change. Most studies of biodiversity change rely on historic data that tend to consist of species records collected by a range of different recorders using different (rarely standardized) methods over extensive time periods. This may result in differences in quality between historic and current datasets. To address this problem, we performed our analyses using different threshold levels of data acceptability (i.e. numbers of records needed per observation unit for inclusion in the analysis), and different periods of data accumulation.

To our knowledge, this is the first comprehensive attempt to deal with dynamic aspects of biodiversity over multiple spatial scales.

Material and methods

The data sets

We obtained the British hoverfly data from the National Biodiversity Network (publicly available at <www.searchnbn.net>) which is largely based on the Hoverfly Recording Scheme (Ball and Morris 2000). For The Netherlands we used data held by the European Invertebrate Survey, which includes data from the Netherlands Syrphidae recording scheme which is a collaboration between the European Invertebrate Survey – the Netherlands, the Netherlands Youth Federation for the study of Nature (NJN) and the Diptera division of the Netherlands Entomological Society (Reemer et al. 2009). Both UK and NL data consist of field observations made by voluntary recorders and of data extracted from collections and literature. Each database record corresponds to one identified individual. Numbers of records in each of the data sets are given in Table 1. We adjusted taxonomic nomenclature in both databases according to the Syrph the Net database (Speight and Castella 2006).

We are explicit about the fact that even such extensive datasets have pitfalls when it comes to estimating temporal

Table 1. Time periods used to split the data and number of records in each period.

Time periods	Number of records – NL	Number of records – UK
1900–1979/1980–2005	78 860/316 042	51 965/294 609
1954–1979/1980–2005	61 938/316 042	37 814/294 609

biodiversity trends. First, there are fewer records from the past in both countries (Table 1, Supplementary material Fig. S1, S4). We controlled for this inequality by using rarefaction (see below). Second, we detected slight temporal increases in the of proportion of common species in the data (Supplementary material Fig. S2). This may be caused either by change in behaviour of recorders, or it can be a real trend. We argue that if this bias is due to recorders, than it would itself (at least in the Netherlands) lead to the opposite trend to what we report here. Third, even if the recorders had not changed their behaviour, this kind of data would always tend to under-represent common species because some amateur recorders might focus on rare and "attractive" species. However, if this kind of bias did not change over time, the comparison of temporal change in biodiversity would be legitimate (Gryntes and Romdal 2008).

Estimation of species richness change

To examine biodiversity change, we contrasted two time periods. We split the datasets to create pre-1980 and post-1980 (including 1980) record sets. Although splitting the data at 1980 is arbitrary, it allowed us to compare the results with previous work (Biesmeijer et al. 2006). Both of our UK and the Dutch data sets span from ~1900 to 2005; if we split them to pre- and post-1980, then the two periods are of unequal lengths (ca 79 and 25 yr). This may have advantages, as it can partially compensate for the generally lower recording effort in the past. However, it may be argued that the use of unequal sampling periods could bias our results because of the species–time relationship (White 2007) – the increase in species numbers at a site with increasing monitoring period. However, discarding all data from the first half of the 20th century would limit our temporal scope and restrict the analysis to relatively short-term changes. Hence, we did our analyses for two time splits: 1) 1900–1979/1980–2005 (unequal time periods as used in Biesmeijer et al. 2006 and 2) 1954–1979/1980–2005 (equal time periods as used in Hickling et al. 2006).

Pre-1980 and post-1980 record sets were not directly comparable because of the uneven number of records. We accounted for that separately within each grid cell by performing individual-based rarefaction (Gotelli and Colwell 2001) on the data from the period with higher number of records, allowing us to "rarefy" down to the number of records in the other survey period. We used one database record as an equivalent of one "individual" (see above and Gotelli and Colwell 2001). We then expressed the temporal change of species richness (ΔS) in each grid cell as

$$\Delta S = (S_{post980} - S_{pre1980})/S_{pre1980} \times 100 \qquad (1)$$

where $S_{post1980}$ and $S_{pre1980}$ are values of species richness in the period before and after 1980 respectively (one of the S values is always obtained through rarefaction). We calculated the ±95% confidence limits of ΔS within each grid-cell using the formulas by Heck et al. (1975) (see also Biesmeijer et al. 2006).

To show the overall trend of ΔS in a set of grid cells we adopted a meta-analytical approach and calculated the weighted bootstrapped median of the ΔS values and ±95% confidence intervals of the bootstrapped median (10 000 resamplings) (Gurevitch and Hedges 1999). We used the median rather than mean because distributions of ΔS were sometimes skewed. We weighted values of ΔS by the inverse of the span of the 95% confidence limits of ΔS. This put higher weights onto grid cells with relatively high numbers of records in both periods and/or with equivalent numbers of records in the two periods, whereas grid cells with a low number of records in one of the periods and/or very uneven sampling efforts were down-weighted. The procedure allowed for a more powerful analysis of species richness change than previously used in Biesmeijer et al. (2006), who used the rather conservative "vote counting" approach (Gurevitch and Hedges 1999), i.e. a comparison of the number of cells where diversity has increased/not changed/decreased.

When working with gridded biodiversity data there is always the problem of spatial dependence of data (spatial autocorrelation; Fortin and Dale 2005). We measured the degree of spatial autocorrelation at the first distance class (correlation of ΔS values in two adjacent grid cells) by calculating the Moran's I (Fortin and Dale 2005) for data with the minimum number of 100 records per grid cell (see below for analysis of the role of recording level). We also tested significance of Moran's I by permutation test (199 permutations).

The role of spatial scale

To investigate whether spatial scale affects the observed patterns of ΔS, we carried out analyses using 10 × 10, 20 × 20, 40 × 40, 80 × 80 and 160 × 160 km grid resolutions. Rather than using a regular grid, we positioned the grid cells so that we obtained as many cells as possible at each spatial scale. We excluded all grid cells with <75% of their area covered by grid cells of the 10 × 10 scale (in most cases, cells over the sea). Sometimes the grid cells overlapped the sea but it was still possible to compensate the lost of mainland area by shifting or elongating of the grid cells to other directions. In such cases we attempted to keep the shape as closer to a square as possible. While greatly elongated sampling units are expected to increase species richness somewhat (Kunin 1997), the small deviations in shape used here would be expected to have negligible effects, which should be more than offset by the advantage of making cell land area more equal than would have otherwise been possible.

When using this area-based definition of scales the extent of the data is not the same at all scales because different sets of grid cells meet our criteria. To make sure that the observed patterns of ΔS are robust to this we conducted a supplementary "agglomerative" analysis in which the set of focal samples remained constant across scales of analysis. We began with the set of high quality 10 km grid cells (>100 records, not more than 10 fold difference in the number of records in the two time periods and the ratio of records/number of species higher than 1.5). After comparing change in that sample itself (level 1), we combined pairs of nearest-neighbouring cells, and repeated the analysis (level 2). We continued by combining neighbouring pairs to create sets of 4 grid cells (level 3) and so on until we grouped the whole set of grid cells together into a single one (level n). For each of these levels we evaluated the ΔS by the same way as described in the previous section.

The role of data quality

Hoverflies are among the best recorded invertebrates, and the UK and the NL are among the best-studied regions on Earth; nonetheless, biodiversity records are patchy and incomplete even in these countries. At the same time, it is unclear what constitutes a sufficient number of records for an area to allow for adequate biodiversity comparisons. A low cut-off (i.e. including cells with few records) would ensure the inclusion of many grid cells in the analysis, but does so by including sites of poorer data quality and possibly lower comparability; conversely a high cut-off results in more robust results, but for few sites only. To test for such trade-offs, we performed our analyses using a range of thresholds for available records per grid cell: 10, 25, 50, 75, 100, 125, 150, 200, 250 and 300 records. Additional inclusion criteria were: 1) a records-to-species ratio of at least 1.5 in each of the two time periods and 2) a less than 10-fold difference in numbers of records between the two time periods (following Biesmeijer et al. 2006). Note that the actual records-to-species ratios were generally much better than 1.5 (Supplementary material Fig. S3).

To summarize, we performed the analysis for all combinations of the following factors: two time splits (1900–1979/1980–2005 and 1954–1979/1980–2005), 10 levels of minimum number of records per grid cell and 5 spatial scales. Data processing and analyses were done in R (R Development Core Team 2009), spatial autocorrelation was measured in SAM 3.0 (Rangel et al. 2006) and maps were created in ArcGIS 9.2 (ESRI).

Results

The NL data consisted of about an order of magnitude more records per grid cell than the UK data (Supplementary material Fig. S1). In addition, in each country the post-1980 period generally contained almost an order of magnitude more records per grid cell than the pre-1980 period at all spatial scales (Supplementary material Fig. S1). Numbers of grid cells that passed our data quality criteria and that were used in the bootstrap analyses are given in Supplementary material Table S1.

Our most important result is that the medians of the observed ΔS were significantly scale dependent (Fig. 2, 3). The 95% confidence intervals (Fig. 2, 3) clearly show that medians of ΔS were at some scales significantly different from zero (above or below) and at some scales indistinguishable from zero. The nature of this dependence varied between the two studied countries. In the Netherlands, when we used the uneven time periods (1900–1979/1980–2005) we did not observe a strong signature of scale dependence in ΔS (Fig. 2, 3), with no significant shift in richness at any scale. Using the equal time periods (1954–1979/1980–2005), ΔS was found to change with spatial scale, with a significant increase in species richness ($\Delta S > 0$) found at finer scales (10 × 10 and 20 × 20 km) (Fig. 2, 3). From the 40 × 40 km scale and above, there was no significant change of species richness ($\Delta S \approx 0$).

In the United Kingdom, ΔS showed some evidence of positive scale-dependence; however, this result was again sensitive to the length of the time periods used (Fig. 2, 3). Using the uneven time periods, species richness decreased at finer scales ($\Delta S < 0$; 10 × 10 and 20 × 20 km) and at the coarsest scale (160 × 160 km). At intermediate scales (40 × 40 and 80 × 80 km) the values of ΔS were not significantly different from 0 ($\Delta S \approx 0$; Fig. 2, 3). When looking at the equal time periods we detected a significant decrease in species richness ($\Delta S < 0$) at 10 × 10 km scale, but this pattern depended on the minimum number of records. There was no significant change in species richness at 20 × 20 and 40 × 40 scales, but a significant increase ($\Delta S > 0$) was shown at 80 × 80 and 160 × 160 km scales. In both UK and the NL the variation of ΔS was higher at finer spatial scales, decreasing towards coarse scales (Fig. 4, 5).

The values of ΔS in the Netherlands were robust to shifting of the data quality criterion (x-axis at Fig. 2) except for sensitivity at the 10 × 10 km scale. The UK results were more sensitive to the number of records, an effect that was most apparent at the 10 × 10 and 20 × 20 scales (Fig. 2). The detected values of ΔS in both countries were also dependent upon whether we compared even or uneven time periods (grey versus black symbols in Fig. 2). Excluding the pre-1954 data generally shifted ΔS towards positive values (Fig. 2, 3).

When we calculated ΔS only over the set of high quality 10 × 10 km grid cells and kept the spatial extent strictly constant at all scales (using an agglomerative approach by merging pairs of nearest-neighbour samples) we no longer detected any sign of scale-dependent ΔS in the Netherlands (Fig. 3). However, in the UK the ΔS was still scale-dependent (Fig. 3), showing a significant decrease in species richness at fine scales (aggregation levels 1 and 2), shifting to essentially zero change at medium to coarse scales (levels 3 to 6), at least where time periods were equal. As in the equal-area approach, the use of unequal time intervals generally resulted in more negative estimates of ΔS.

We did not detect significant spatial autocorrelation in ΔS values in either the UK or NL apart from very weak autocorrelation at 10 × 10 scale in NL (Moran's I values in Fig. 4, 5). However, the significance tests at 80 × 80 and 160 × 160 km scales may lack power due to low number of grid cells at these scales.

Discussion

Spatial scale and biodiversity change

We found that rates of species richness change are different at different spatial scales. Interestingly, while ΔS shifted with scale in both countries, the directions of change differed. In the Netherlands, we document fine scale species richness gains that disappear at coarser scales, which suggests a decrease in species turnover across space (with lower β diversity reducing the slope of the SAR, as illustrated in Fig. 1b, Supplementary material Fig. S1). This suggests that the increase in fine scale species richness in NL hoverfly assemblages has been accompanied by a

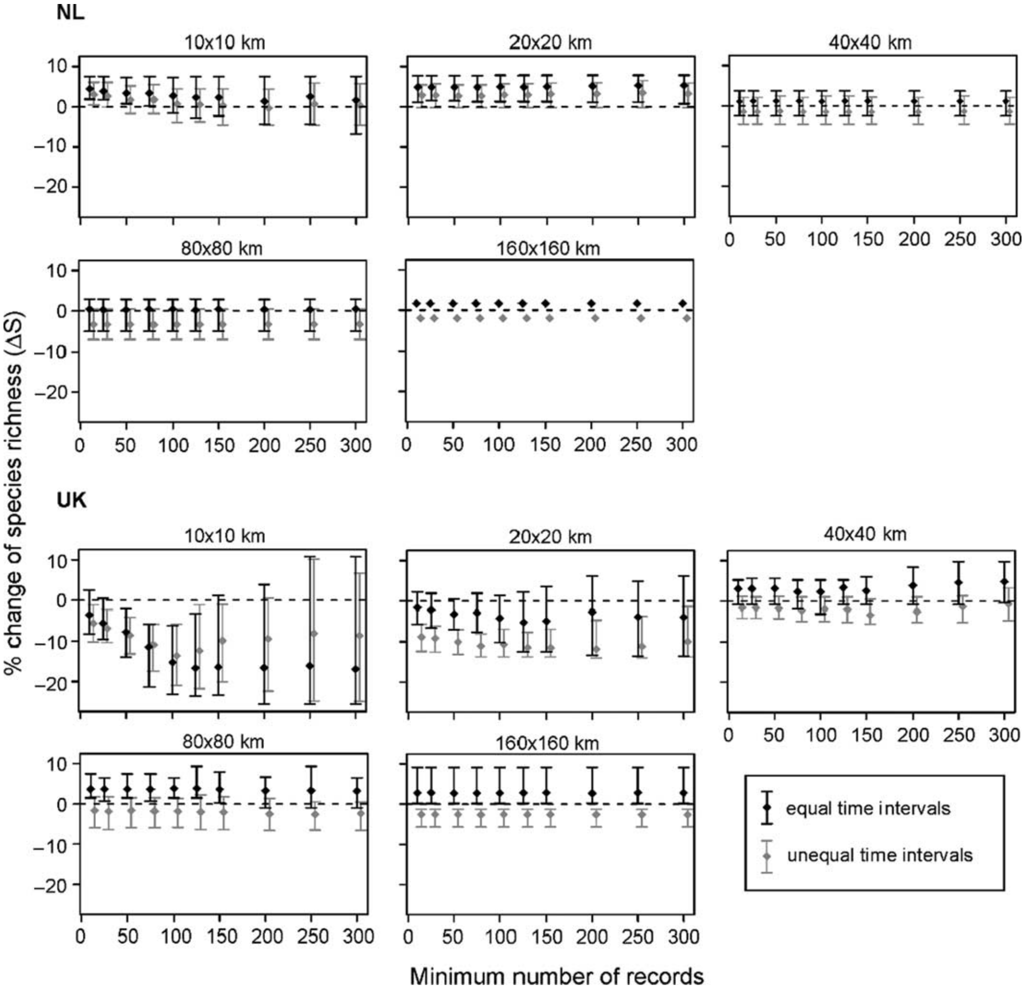


Figure 2. Species richness changes (ΔS) in the Netherlands and the UK between two pre- and post-1980 periods and their dependence on spatial scale and minimum number of records criteria (data quality). Diamonds are bootstrapped medians (10 000 resamplings). The bars represent 95% bootstrapped confidence intervals of the medians. Comparisons based on different pre-1980 periods are indicated by shading, with black for equal intervals (1954–1979 and 1980–2005), and grey for analyses using unequal intervals (1900–1979 and 1980–2005). The dashed line represents zero change.

degree of biotic homogenisation (McKinney and Lockwood 1999), which may reflect the growing dominance of a few abundant and/or strongly dispersing species (cf. McKinney and LaSorte 2007, Supplementary material Fig. S2). Conversely, in the UK there were fine scale losses in richness shifting to coarse scale gains, indicating an increase in β diversity (Fig. 1d, Supplementary material Fig. S1). Such shifts might be expected during times of biotic transition, when substantial fine scale losses of native species have not yet resulted in coarse-scale losses (Thomas and Abery 1995). More generally, the observed shifts in β diversity must reflect changes in the level of aggregation of species' distributions or changing the sizes of distributions, which may in turn reflect one of a number of environmental drivers (increase/decrease of landscape heterogeneity or dispersal barriers; Storch et al. 2008). Given the different scaling properties of different potential drivers of change (see Introduction), it is perhaps not surprising that no single general pattern of change across scales emerges from our results for the two countries considered.

An alternative approach, based on the sample agglomeration method showed generally weaker trends overall, in part because of greater variation around median effects – perhaps due to variation in how far apart nearest neighbours were. While this variation obscured any scale dependence in NL change patterns, the strong signal of negative species richness change at fine scales in the UK, disappearing at coarser scales is maintained. Thus even this conservative approach provides strong evidence of scale specificity in biodiversity change.

Our findings of scale-specific shifts in hoverfly richness are supported by comparing different (scale specific) studies on other taxa. Thus for example, UK butterfly species richness has apparently increased 7.62% at 20 × 20 km scale (Menendez et al. 2006), but it has decreased 5% nationally (Konvicka et al. 2006), which mirrors our findings for NL (but not UK) hoverflies. Conversely, recent studies of UK plants have found substantial declines in diversity at very fine spatial scales (Haines-Young et al. 2003), while there are apparent increases in diversity at coarser scales (Preston et al. 2002), although the latter is likely due (at least in part) to increased recorder effort.

The fact that species richness change is scale-dependent has implications for both basic biodiversity research and for

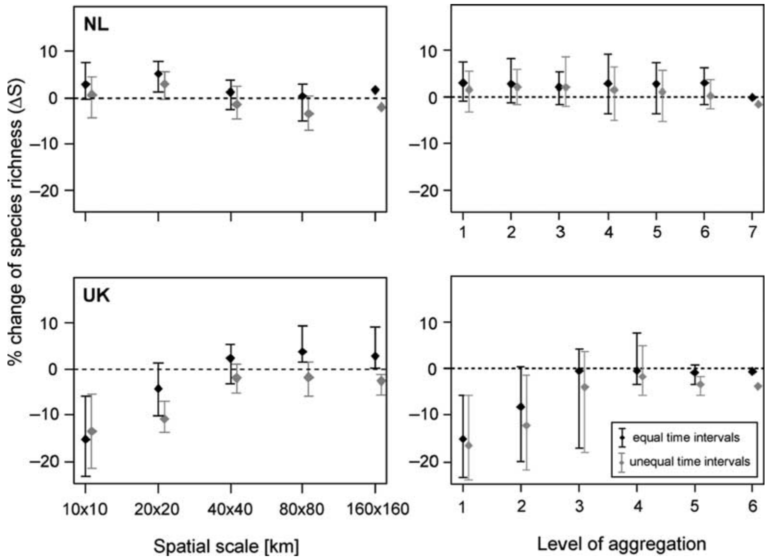


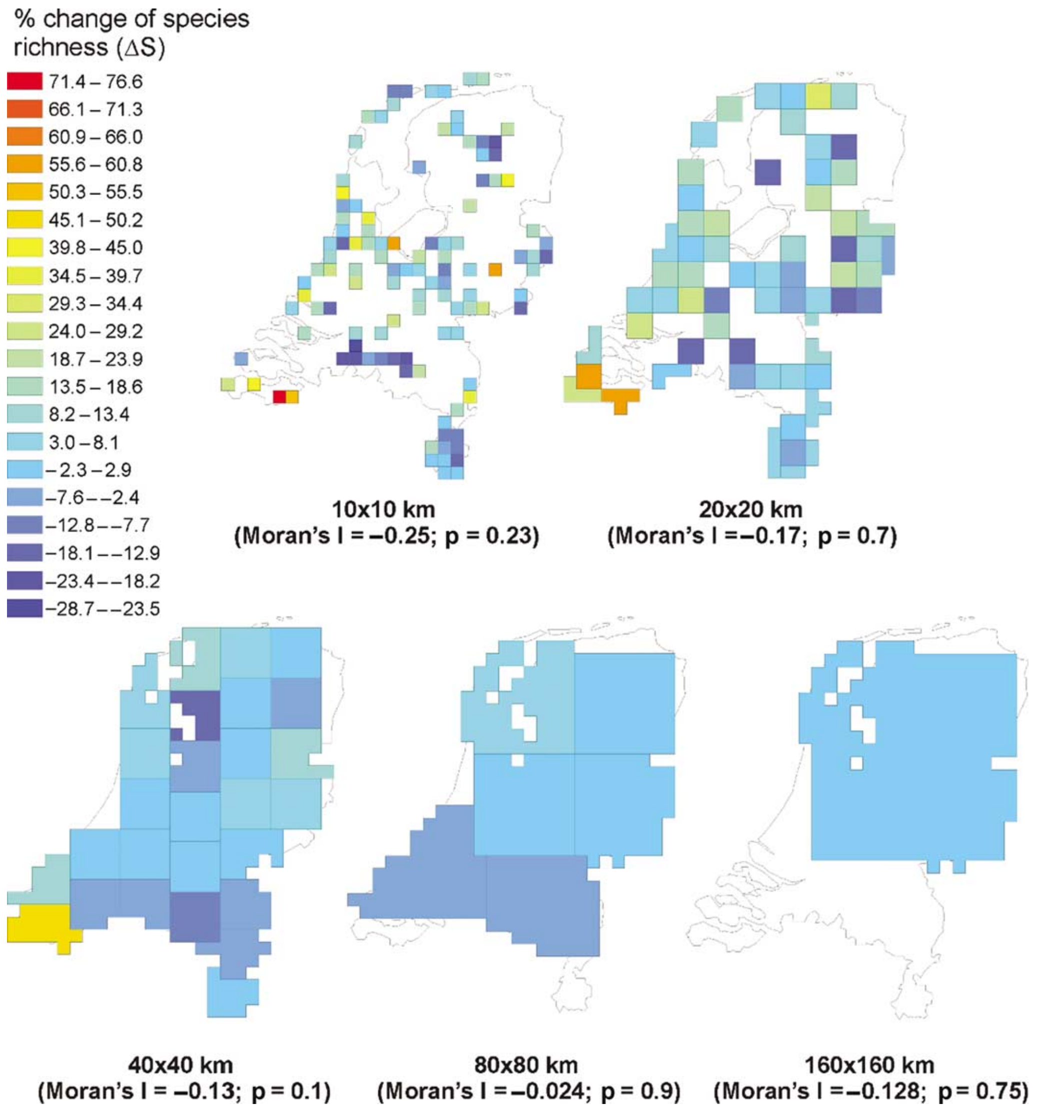
Figure 3. Species richness changes (ΔS) in the Netherlands and the UK between two pre- and post-1980 periods and their dependence on spatial scale (left) and on the level of aggregation when the extent of the analyses was kept strictly constant (right; see Material and methods for the description of the aggregation algorhitm). Here we show ΔS only for grid cells with >100 records, not more than 10 fold difference in the number of records in the two time periods and the ratio of records/number of species higher than 1.5. Comparisons based on different pre-1980 periods are indicated by shading, with black for equal intervals (1954–1979 and 1980–2005), and grey for analyses using unequal intervals (1900–1979 and 1980–2005). The dashed line represents zero change.

applied ecology. Most of the high-profile studies reporting past species richness changes (Lemoine and Bohning-Gaese 2003, Biesmeijer et al. 2006, Konvicka et al. 2006, Menendez et al. 2006) or predicting future scenarios (Thuiller et al. 2005, Jetz et al. 2007) were performed at single spatial scale. The scales used in these studies fall within the range covered in our study, raising concerns that the scale-sensitivity of dynamics documented here might apply to them as well. Our results suggest that any study that operates only at a single scale provides only a partial picture of biodiversity change, and should be interpreted with explicit reference to the particular scale of observation.

Assessments of biodiversity change are essential in the evaluation of conservation decisions (e.g. European Agri-envionment Schemes) and monitoring of progress towards biodiversity targets (e.g. CBD target of halting biodiversity loss by 2010). Our results indicate that such evaluations should not be restricted to a single spatial scale. At the very least, we need to develop cross scale methods of assessing biodiversity change (perhaps building on those employed here), to allow the upscaling of local biodiversity data to monitor coarser scale aspects of biotic change. Moreover, there are good reasons to expect scale specific effects of management on biodiversity. For example, the introduction of an agri-environmental scheme uniformly across a country might be expected to make habitats more similar across space, potentially reducing β diversity, even if it increases diversity at a local scale (α diversity). In contrast, heterogeneity or local decision-making in the introduction or nature of agri-environmental schemes might enhance both α and β diversity. Knowledge of the relationship between species diversity and spatial scale will also allow for the design of biodiversity management schemes that maintain or even enhance local and regional biotic differences, and thus enhance biodiversity on many scales.

Spatial autocorrelation and spatial averaging

The absence (or near absence) of spatial autocorrelation at most spatial scales we analyzed is remarkable (the only exception is the very weak autocorrelation at 10 × 10 km scale in the NL). Many adjacent grid cells show opposite trends of biodiversity change (Fig. 4, 5). This means that current biodiversity dynamics are probably not governed predominantly by a single, globally acting factor such as climate. Instead, it appears that site-specific factors (e.g. local land use change) most strongly influence biodiversity dynamics of hoverflies in our sample, or alternatively that the signal of climate is blurred by stronger effect of environmental noise or stochastic recorder behaviour. If change in species richness is governed at fine scales and is not spatially autocorrelated than it should average out as we move from fine to coarse scales and species richness changes should show stronger relative fluctuations at fine scales. This is exactly what we observed in both the UK and NL (Fig. 4, 5). Scale dependent dynamics of species richness change would than be a consequence of spatial averaging (Levin 1992). Finally, the changes we documented are surprisingly large at some grid cells (Fig. 4, 5). In several areas more than a third of the species recorded have disappeared while an equal proportion has colonized an adjacent cell. Though this may be a real image of very dynamic local-scale processes (e.g. conversion to agriculture, loss of flower-rich grasslands, afforestation), we cannot exclude to possibility that it reflects an effect of sampling (in spite of the efforts undertaken to control for this and using some of the best large-scale biodiversity datasets available).



10x10 km (Moran's I = –0.25; p = 0.23)

20x20 km (Moran's I = –0.17; p = 0.7)

40x40 km (Moran's I = –0.13; p = 0.1)

80x80 km (Moran's I = –0.024; p = 0.9)

160x160 km (Moran's I = –0.128; p = 0.75)

Figure 4. Patterns of species richness change (ΔS) in the Netherlands between the 1954–1979 and 1980–2005 periods. ΔS was calculated only for grid cells that met the data quality criteria (>100 records; not more than 10 fold difference in the number of records in the two time periods; the ratio of records/number of species higher than 1.5).

Data quality affects pattern of biodiversity change

Our analyses rely on biodiversity data collected in an unsystematic way by amateurs and professional entomologists. It is therefore important that the biases and limitations in data of this sort be recognised (see also above). Even though the hoverfly data used here are among the best in the world for any invertebrate taxon or region, we find that data quality (e.g. recorder effort and coverage) influences the outcomes of our analyses of biodiversity change. Our analyses reveal striking issues related to spatial and temporal data quality.

Spatial data quality

The UK records are aggregated in fewer areas with high sampling effort than the NL records (Supplementary material Fig. S4). Also, the data from the UK are spread more thinly (average number of records 1 km^{-2} in the pre-1980 period is 2.33 in NL and only 0.21 in UK). Therefore, results from the UK should be interpreted with some caution.

As noted in the introduction, there is an intrinsic trade-off in setting data-quality thresholds for studying biodiversity trends in species records: a low threshold allows large numbers of sites to be examined, but with low confidence in individual results; whilst a high quality threshold provides greater confidence in individual results, but allows fewer sites to be assessed. Perhaps surprisingly, our results proved relatively robust over a wide range of quality thresholds. In the Netherlands, this may be because a large fraction of sites had high enough data quality to be included whatever the threshold, while in the UK (with lower record density overall) some trends were obscured at low thresholds or disappeared at high thresholds. Overall, a requirement for 50–100 records per 10 × 10 km cell (cf. Biesmeijer et al. 2006) seems to perform well, providing good resolution of trends while retaining substantial site numbers. Our use of the tools of meta-analysis to combine results across sites allows even lower thresholds to be considered; in general,

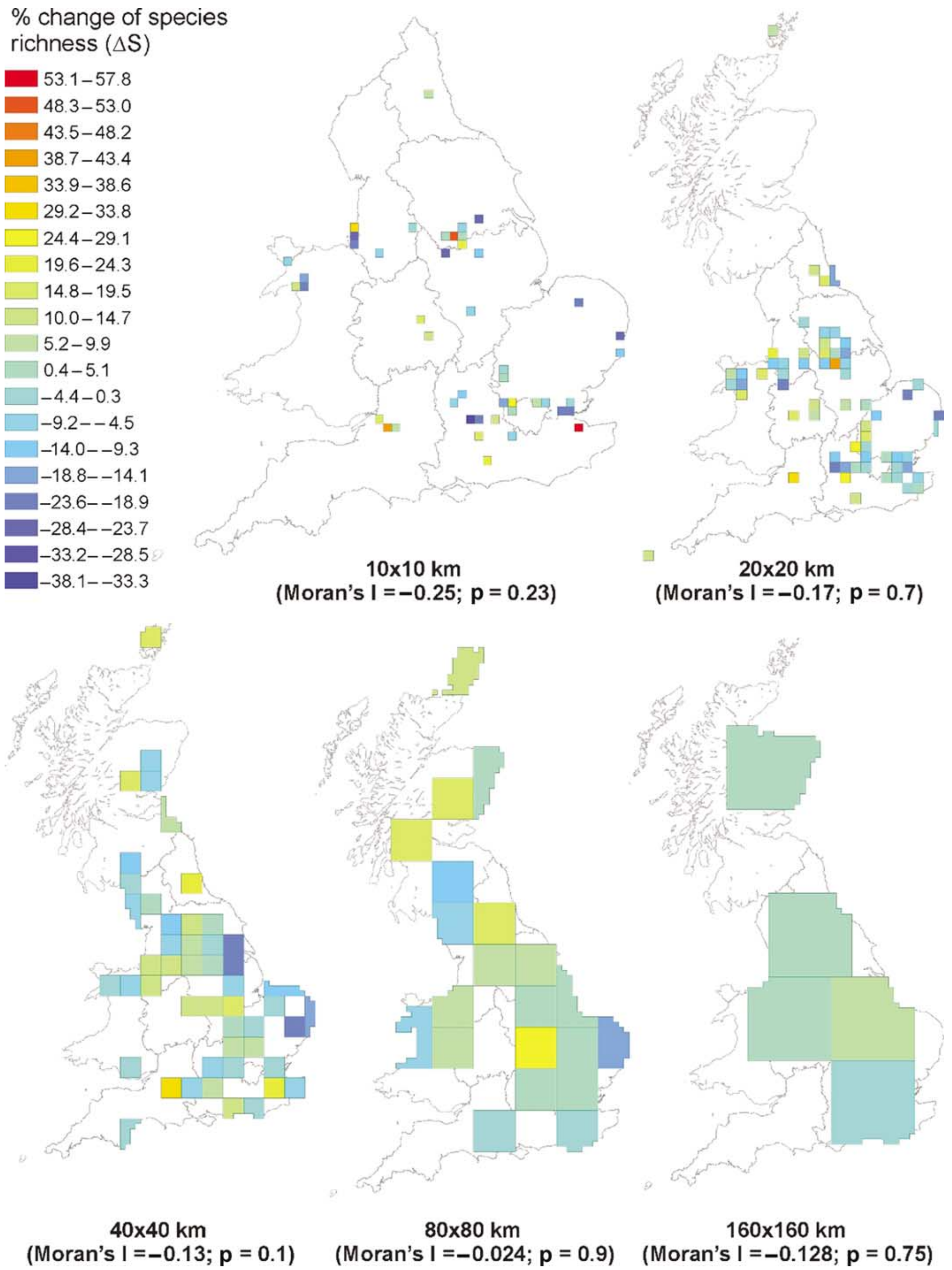


Figure 5. Patterns of species richness change (ΔS) in the United Kingdom between the 1954–1979 and 1980–2005 periods. ΔS was calculated only for grid cells that met the data quality criteria (see caption of Fig. 4).

confidence intervals became narrower at lower quality cut-offs. Too high a threshold had real costs on precision, with expanding confidence intervals obscuring most trends.

Temporal data quality

Apart from the incomplete spatial coverage of sampling effort there is the fact that in all analyses post-1980 data had more records than the pre-1980 data (when the records were summed across all grid cells). Can our results be explained by this bias? The first step to avoid this is the rarefaction procedure itself which is robust to differences in sample size (Gotelli and Colwell 2001). We also showed that the observed rate of species richness change can be sensitive to the inequality of time periods. When the pre-1980 period was longer than the post-1980 period, the observed trend in biodiversity was virtually always made more negative. Sometimes the effects were strong enough to cause a shift in perceived diversity trends; e.g. in the 160 × 160 km scale UK results, ΔS was positive when equal periods were compared, but significantly negative when a longer pre-1980 period was used. This can be explained by the species-time relationship (White 2007), whereby the total number of species found in an area grows as the time period considered gets longer, in part due to integrating across a wider range of climatic and environmental conditions. The use of a longer pre-1980 period thus artefactually expands the perceived species richness of this initial period, and thus creates a bias towards perceiving declines in species richness. Our results stress the importance of keeping the time periods equal in any study dealing with biodiversity dynamics.

Future prospects

In this paper we have developed a simple framework to analyze multi-scale patterns of biodiversity change using data suffering from unequal and haphazard sampling efforts. Such methods are essential, as only a tiny fraction of "charismatic" taxa (in particular, birds butterflies and vascular plants) have been subject to systematic monitoring programmes or studied well enough for full distributional atlases to be prepared; and only in a few western European countries. Scattered and unsystematic point records are thus all that is available for the vast majority of the Earth's species. If the preservation of biodiversity is a global and regional priority (as enshrined in the CBD and Göteborg targets), then we need to develop tools to allow meaningful estimates of biodiversity change for such taxa; we hope this paper will make a contribution in this regard. Indeed, the multi-scale analyses explored here might well be profitably adapted even to well-studied groups such as butterflies, birds or plants, as little explicit information is available to date about their dynamics at multiple scales.

Acknowledgements – We thank to David Storch, Francis Gilbert, Daniel J. McGlinn and one anonymous referee for valuable comments. We are indebted to the many individuals who have created and participated on the British Hoverfly Recording Scheme and the Netherlands Syrphidae recording scheme. PK was supported by the Marie Curie Early Stage Training fellowship at Univ. of Leeds (UK) and by the Czech Ministry of Education (LC06073). The study was also supported by the FP6 EU project ALARM (COCE-CF-2003-506675) and by the FP7 EU project SCALES (contract number 226852).

References

Arrhenius, O. 1921. Species and area. – J. Ecol. 9: 95–99.

Ball, S. G. and Morris, R. K. A. 2000. Provisional atlas of British hoverflies. – Centre for Ecology and Hydrology.

Biesmeijer, J. C. et al. 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. – Science 313: 351–354.

Chase, J. M. and Leibold, M. A. 2002. Spatial scale dictates the productivity–biodiversity relationship. – Nature 416: 427–430.

Davies, K. F. et al. 2005. Spatial heterogeneity explains the scale dependence of the native–exotic diversity relationship. – Ecology 86: 1602–1610.

Diniz-Filho, J. A. F. et al. 2003. Spatial autocorrelation and red herrings in geographical ecology. – Global Ecol. Biogeogr. 12: 53–64.

Englund, G. and Hamback, P. A. 2007. Scale dependence of immigration rates: models, metrics and data. – J. Anim. Ecol. 76: 30–35.

Fortin, M. and Dale, M. 2005. Spatial analysis: a guide for ecologists. – Cambridge Univ. Press.

Gotelli, N. J. and Colwell, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. – Ecol. Lett. 4: 379–391.

Gryntes, J. A. and Romdal, T. S. 2008. Using museum collections to estimate diversity patterns along geographical gradients. – Folia Geobot. 43: 357–359.

Gunton, R. M. and Kunin, W. E. 2007. Density effects at multiple scales in an experimental plant population. – J. Ecol. 95: 435–445.

Gunton, R. M. and Kunin, W. E. 2009. Density-dependence at multiple scales in experimental and natural plant populations. – J. Ecol. 97: 567–580.

Gurevitch, J. and Hedges, L. V. 1999. Statistical issues in ecological meta-analyses. – Ecology 80: 1142–1149.

Haines-Young, R. et al. 2003. Changing landscapes, habitats and vegetation diversity across Great Britain. – J. Environ. Manage. 67: 267–281.

Hanski, I. 1999. Metapopulation ecology. – Oxford Univ. Press.

Hartley, S. and Kunin, W. E. 2003. Scale dependency of rarity, extinction risk, and conservation priority. – Conserv. Biol. 17: 1559–1570.

Heck, K. L. et al. 1975. Explicit calculation of rarefaction diversity measurement and determination of sufficient sample size. – Ecology 56: 1459–1461.

Hickling, R. et al. 2006. The distributions of a wide range of taxonomic groups are expanding polewards. – Global Change Biol. 12: 450–455.

Jarosik, V. and Lapchin, L. 2001. An experimental investigation of patterns of parasitism at three spatial scales in an aphid-parasitoid system (Hymenoptera: Aphidiidae). – Eur. J. Entomol. 98: 295–299.

Jetz, W. et al. 2007. Projected impacts of climate and land-use change on the global diversity of birds. – PLoS Biol. 5: 1211–1219.

Konvicka, M. et al. 2006. Butterfly extinctions in European states: do socioeconomic conditions matter more than physical geography? – Global Ecol. Biogeogr. 15: 82–92.

Kunin, W. E. 1997. Sample shape, spatial scale and species counts: implications for reserve design. – Biol. Conserv. 82: 369–377.

Kunin, W. E. 1998. Extrapolating species abundance across spatial scales. – Science 281: 1513–1515.

Lemoine, N. and Bohning-Gaese, K. 2003. Potential impact of global climate change on species richness of long-distance migrants. – Conserv. Biol. 17: 577–586.

Levin, S. A. 1992. The problem of pattern and scale in ecology. – Ecology 73: 1943–1967.

McKinney, M. L. and Lockwood, J. L. 1999. Biotic homogenization: few winners replacing many losers in the next mass extinction. – Trends Ecol. Evol. 14: 450–453.

McKinney, M. L. and La Sorte, F. A. 2007. Invasivenes and homogenization: synergism of wide dispersal and high local abundance. – Global Ecol. Biogeogr. 16: 394–400.

Menendez, R. and Thomas, C. D. 2000. Metapopulation structure depends on spatial scale in the host-specific moth *Wheeleria spilodactylus* (Lepidoptera: Pterophoridae). – J. Anim. Ecol. 69: 935–951.

Menendez, R. et al. 2006. Species richness changes lag behind climate change. – Proc. R. Soc. B 273: 1465–1470.

Pocock, M. J. O. et al. 2006. Ecological correlates of range structure in rare and scarce British plants. – J. Ecol. 94: 581–596.

Preston, C. D. et al. 2002. New atlas of the British and Irish flora: an atlas of the vascular plants of Britain, Ireland, The Isle of Man and the Channel Islands. – Oxford Univ. Press.

R Development Core Team 2009. R: a language and environment for statistical computing. – R Foundation for Statistical Computing.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. – Ecol. Lett. 8: 224–239.

Rahbek, C. and Graves, G. R. 2001. Multiscale assessment of patterns of avian species richness. – Proc. Natl Acad. Sci. USA 98: 4534–4539.

Rangel, T. et al. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. – Global Ecol. Biogeogr. 15: 321–327.

Reemer, M. et al. 2009. De Nederlandse zweefvliegen (Diptera: Syrphidae). – Nederlandse Fauna 8: 1–442.

Rosenzweig, M. L. 2001. The four questions: what does the introduction of exotic species do to diversity? – Evol. Ecol. Res 3: 361–367.

Shmida, A. and Wilson, M. V. 1985. Biological determinants of species diversity. – J. Biogeogr. 12: 1–20.

Speight, M. C. D. and Castella, E. 2006. StN Database: content and glossary of terms, Ferrara, 2006. – In: Speight, M. C. D. et al. (eds), Syrph the Net, the database of European Syrphidae. Syrph the Net publications, Dublin, pp. 1–77.

Stohlgren, T. J. et al. 2003. The rich get richer: patterns of plant invasions in the United States. – Front. Ecol. Environ. 1: 11–14.

Stohlgren, T. J. et al. 2006. Scale and plant invasions: a theory of biotic acceptance. – Preslia 78: 405–426.

Storch, D. et al. (eds) 2007. Scaling biodiversity. – Cambridge Univ. Press.

Storch, D. et al. 2008. The quest for a null model for macroecological patterns: geometry of species distributions at multiple spatial scales. – Ecol. Lett. 11: 771–784.

Thomas, C. D. and Abery, J. C. G. 1995. Estimating rates of butterfly decline from distribution maps – the effect of scale. – Biol. Conserv. 73: 59–65.

Thomas, J. A. et al. 2004. Comparative losses of British butterflies, birds, and plants and the global extinction crisis. – Science 303: 1879–1881.

Thuiller, W. et al. 2005. Climate change threats to plant diversity in Europe. – Proc. Natl Acad. Sci. USA 102: 8245–8250.

Warren, M. S. et al. 2001. Rapid responses of British butterflies to opposing forces of climate and habitat change. – Nature 414: 65–69.

Wiens, J. A. 1989. Spatial scaling in ecology. – Funct. Ecol. 3: 385–397.

White, E. P. 2007. Spatiotemporal scaling of species richness: patterns, processes, and implications. – In: Storch, D. et al. (eds), Scaling biodiversity. Cambridge Univ. Press, pp. 325–346.

Whittaker, R. J. et al. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. – J. Biogeogr. 28: 453–470.

Willis, K. J. and Whittaker, R. J. 2002. Species diversity – scale matters. – Science 295: 1245–1248.

Wilson, R. J. et al. 2004. Spatial patterns in species distributions reveal biodiversity change. – Nature 432: 393–396.