

ORIGINAL RESEARCH

Core and occasional species: A new way forward

Alvin J. Helden^{1,2} 

¹Applied Ecology Research Group, School of Life Sciences, Anglia Ruskin University, Cambridge, UK

²UCD School of Agriculture and Food Science, University College Dublin, Dublin 4, Ireland

Correspondence

Alvin J. Helden, Applied Ecology Research Group, School of Life Sciences, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK.

Email: alvin.helden@anglia.ac.uk

Funding information

Environmental Protection Agency, Ireland, Grant/Award Number: 2001-CD/B1-M1

Abstract

1. Various methods have been used to divide communities into core species and occasional or satellite species. Some methods are somewhat arbitrary, and there is evidence that many communities are more multimodal than bimodal. They also tend to rely on having multiple years of data.
2. A completely novel method is presented that not only has no requirement for long-term datasets but can divide communities into multiple groups. It is based on probability a species is present, calculated using Simpson's index and the sequential removal of species from the data.
3. The sequential Simpson's index method was applied to species data from a grassland insect community. It was also applied to eleven other datasets that had been divided into core and occasional species in previously published studies.
4. The new method was found not only to be consistent with previous core-occasional assessments but also able to identify multimodality in species abundance distributions.
5. Although ideally used with a measure of persistence (frequency of occurrence) to rank species, community structure is consistently described even with only species abundance data.
6. The method can be applied to short or long-term datasets and can help identify multimodality and provide valuable insight into how communities change in time or space.

KEYWORDS

community structure, core and occasional species, grassland insects, multimodality, Simpson's index

1 | INTRODUCTION

The idea of dividing up communities into species groups, based on their relative abundance or frequency of occurrence (persistence), is long standing (e.g., see Winterbottom, 1949). However, it was arguably not approached in a more formalized way until the core-satellite hypothesis of Hanski (Hanski, 1982; Magurran, 2007; Supp

et al., 2015), in which distribution the core species are found at more sites and are relatively abundant, compared with satellite species. The division of species into groups based on the frequency that species are encountered was incorporated into the UK's National Vegetation Classification (NVC), in which plant species were classified into five frequency classes, based on 20% bands (Rodwell, 2006). In the NVC, the two highest frequency bands were given the term,

The funding from the EPA (Ireland) funded only the work at UCD.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

constant, with the other three groupings being, frequent, occasional, and scarce.

Magurran and Henderson (2003) extended the concept by replacing spatial distribution with temporal frequency and not applying any a priori expectation of frequency classification. They used data from a 21-year study of fish in the Bristol Channel, UK. By plotting maximum yearly abundance against number of years, a species was recorded (persistence); they were able to divide the community into core and occasional species. Core species were more persistent and generally more abundant and ecologically associated with estuarine habitats. Occasional species were infrequent, typically less abundant and not considered estuarine species.

The deconstruction of communities into core and occasional groups is valuable as it may enable changes in the abundance of individual species or species groups to be detected (Magurran & Henderson, 2010; Whittaker, 2015). It can also be used to assist in the study of species abundance distributions (SADs) and the testing of species abundance models (Magurran, 2007). Although a number of models have been proposed to describe SADs (Antão et al., 2017; Magurran, 2004), it has become clear that SADs are frequently not adequately described by them, at least in part due to their multimodal nature (Antão et al., 2017; Matthews, Borges, et al., 2014).

Multimodality can be defined as the presence of two or more species groupings within communities (Antão et al., 2017; Matthews, Borges, et al., 2014; Ugland & Gray, 1982). One way of understanding multimodality is to develop alternative models to describe SADs, such as the Gambin model (Matthews, Borregaard, et al., 2014; Ugland et al., 2007; Whittaker, 2015). Another approach is to identify the groupings that generate multiple modes and study their characteristics.

Since the early work of Hanski (1982), and Magurran and Henderson (2003), many studies, using a range of approaches, have deconstructed communities into occasional or satellite, and core species. The communities studied have been varied, covering a wide range of taxa (Astudillo-García et al., 2017; Coyle et al., 2013; van der Gast et al., 2011; Genner et al., 2004; Gray et al., 2005; Ulrich & Ollik, 2004; Ulrich & Zalewski, 2006; Umaña et al., 2017).

Among these studies, Genner et al. (2004) used a modification of Magurran and Henderson's (2003) method to identify core and occasional species of fish in long-term data from the Bristol and English Channels. Plotting abundance against persistence, they modeled the relationship with a third-order polynomial and divided the community at the inflection point into core and occasional species, with the same Bristol Channel community sub-division as Magurran and Henderson (2003). Ulrich and Ollik (2004) classified core species of forest Hymenoptera as being those found in six or more years from eight, while occasional species were found in three or less years. Several other studies have taken a similar, rather arbitrary approach by defining core species at being one end of a measure of frequency and occasional the other end (Coyle et al., 2013; Hansen & Carey, 2015; Ulrich & Zalewski, 2006). Some have used a single frequency point to divide communities into two groups (Barnes et al., 2016; Dolan et al., 2009; Matthews, Borges, et al., 2014), while

others have considered the presence of intermediate species (Boss & Silva, 2014; Supp et al., 2015). Alternative methods have included using species' variance to mean ratio, calculating Simpson's index for subgroups of species and fitting mathematical models to community data (Boisnier et al., 2010; van der Gast et al., 2011; Gray et al., 2005). There has been criticism that many approaches have been arbitrary and a potential source of artifacts, although some studies have tried to use biological characteristics of the communities to help improve the reliability of species groupings (Barlow et al., 2010; Boisnier et al., 2010; Whittaker, 2015). As well as arbitrariness, any attempt to divide a community into core and occasional species presupposes the hypothesis that a binary division best describes a community. However, such a simple a priori approach may not reflect the reality and complexity of many, or indeed most communities, and their arbitrary nature makes replication of methods problematic. Indeed, with multimodality believed to be widespread (Antão et al., 2017; Dornelas & Connolly, 2008), applying a bimodal description to a community risks missing potentially interesting ecological patterns. Therefore, there is a need for an approach to divide communities in a more descriptive, objective, and reproducible way.

Another limitation of some previous methods has been that core-occasional and core-satellite distinctions have relied on having very large datasets, only available after sampling for many years or at many sites, thus precluding their use from shorter term or smaller scale studies. For example, Magurran and Henderson (2003) was based on 21 years of sampling fish communities. Deconstructing communities without such runs of data is more difficult, and unfortunately, much community data from field research are short-term (Barlow et al., 2010). Finding a reliable method for objectively identifying species groupings without long-term/multiple site sampling would therefore be very valuable. It would allow many more community datasets to be studied in terms of species groupings and would have the additional benefit of making it easier to look at how SADs and changes to species groupings vary over time or space (Magurran, 2007; Magurran & Henderson, 2010).

I studied the grassland insect community from an intensively cattle-grazed site at Teagasc Grange, Ireland, over 4 years, to which I attempted to apply a binary core-occasional partitioning of species. After an initial use of a modified form of the Genner et al. (2004) approach, in which I identified the method of calculating persistence as being potentially problematic, I developed a novel approach based on calculation of the Simpson's index and the sequential removal of species from the data. The new method has the advantage over previous partitioning methods of being descriptive and making no a priori assumption regarding the pattern of species groupings within a community, thus moving away from a simple binary division on communities. It can also be applied to a much wider range of datasets, as it is much less reliant on having data from multiple years or sites. The method was tested against data from Magurran and Henderson (2003) and Genner et al. (2004), as well as nine other datasets from previous studies. In this paper, I describe this new method and discuss its application to the study of community structure. Specifically, I test the following questions:

1. Is it sensitive to the method of calculating persistence (frequency of occurrence)?
2. Does it show the same pattern a species grouping when species are ranked by persistence or by abundance?
3. Can it be applied to a range of previous community datasets?

2 | MATERIAL AND METHODS

2.1 | Insect sampling

Coleoptera and Hemiptera were suction sampled ($n = 692$) from the vegetation and ground surface of a cattle-grazed agricultural grassland (Drennan & McGee, 2009; Helden et al., 2015) at Teagasc Grange, Co Meath, Ireland, between 2002 and 2005. Further details of sampling and species identification can be found in Helden et al. (2015).

2.2 | Statistical modeling and index calculations

The statistical modeling and calculation of values for Simpson's index, described below, were carried out using R version 3.4.0 (R Core Team, 2017). Code used for data manipulation and subsequent calculation of Simpson's index is given in Appendix A.1.

2.3 | Application of Genner et al. method to Grange Hemiptera and Coleoptera data

The Genner et al. (2004) method was applied to Grange Hemiptera and Coleoptera species and abundance data, to divide the species into core and occasional species (i.e., a binary classification). With only 4 years of data, year could not be used as a useful measure of persistence. Instead, it was measured by the number of times a species was found in a group of samples. To take account of phenological change, samples from all years and dates were randomly allocated in approximately equal-sized groups. Group size was defined by the number of individuals rather than number of samples. The minimum group size was taken as the number of individuals required to give a stabilized value (i.e., groups of the same or larger size would show no difference) of Simpson's index ($1/D$) (Lande et al., 2000; Magurran, 2004). For Hemiptera, this was a mean group size of 43.6 individuals (115 groups) and for Coleoptera 99.7 individuals (63 groups). For each species, persistence was determined as the number of groups in which it was present.

\log_{10} mean abundance for each species was plotted as the response and persistence as the explanatory variable to give a sigmoidal pattern. A third-order polynomial line ($y = x + x^2 + x^3$) was fitted, and the value for persistence at the inflection point was determined. Species were classified as core if more frequent than the persistence value observed at the inflection point and as occasional if less so.

Although minimum group size could be determined using Simpson's index (see above), it was not known what the effect of larger group sizes would be on the number of core species. Therefore, the procedure was repeated for a series of nine other group sizes for both Coleoptera and Hemiptera. For Coleoptera, these were 146.1 (mean) individuals (63 samples); 202.7 (31); 251.3 (25); 299.2 (21); 349.1 (18); 418.9 (15), 448.8 (14), 523.6 (12), and 628.3 (10). The equivalent for Hemiptera was 94.5 (53); 147.3 (34); 200.4 (25); 250.5 (20); 313.1 (16); 357.8 (14); 417.4 (12); 455.4 (11); 500.9 (10).

2.4 | Proposed new sequential Simpson's index method applied to Grange insect data and to Bristol and English Channel fish data

The proposed new sequential Simpson's index method was applied to Hemiptera and Coleoptera data from Grange, as well as to the Bristol Channel and English Channel fish data from Genner et al. (2004).

For Grange, the minimum group size (as defined in the previous section, above) was used for Hemiptera (115 groups, mean abundance 43.6) and Coleoptera (63 groups, mean 99.7). The frequency that a species was found in the groups was used only to rank species in terms of persistence. Species of equal persistence were ordered by abundance. Simpson's index ($1/D$) was calculated for the full community after adding one to the abundance of each species. Then, the highest ranked species (most persistent) was removed and Simpson's index recalculated. This was repeated with the sequential removal of the most persistent until only one species remained, when Simpson's $1/D = 1$.

Simpson's index was plotted as the response variable with species rank (persistence) as the explanatory variable. The resultant data were modeled with polynomial generalized linear models using Gaussian error structure. A series of models, beginning with $1/D + (1/D)^2$, were created. For each new model, the next sequential power was added, such that the next was a third order polynomial, followed by a fourth order polynomial, and so on. The proportion of deviance explained by each model was calculated using the Dsquared function in the modEvA package (Barbosa et al., 2015). Increasing the polynomial power led to an increase in deviance explained, up to a point at which the deviance showed little further increase. The first model as deviance plateaued was taken as the optimal model. Model simplification was then carried out by the sequential removal of any nonsignificant terms, with adjacent models compared with AIC values and deletion tests used to justify changes (Crawley, 2007).

Polynomial models showed linear alternation in concavity at inflection points. These were identified by solving the model for each species rank and determining between which ranks the model concavity changed (Appendix A.1). The inflection points were taken as indicating dividing points between species groups. Original abundance and persistence data were compared with the inflection points. In some cases, at the less abundant and infrequent end of the data, species groups were combined when abundance and/or persistence were identical or

very similar. In some cases when adjacent species near group boundaries had identical persistence, the boundary of the grouping was moved slightly away from the inflection point itself. The multiple species groupings form a continuum from the most persistent at one extreme to the least persistent at the other. The data were also inspected to try to determine a suitable dividing point between core (C) and occasional (O) species groups. This was judged in part by the position of the inflection points and partly by the abundance and persistence of the species. In most datasets, this was either at the central inflection point, in the case of an even number of groupings, or one of the two central inflection points for an odd number.

To test whether the number of core species defined by the new method was sensitive to group size, it was tested using the same 10 groups of Grange Coleoptera and 10 groups of Hemiptera data, as described in the previous section. This enabled the method to be compared with the Genner et al. (2004) method in terms of group size effect.

The sequential Simpson's method was then applied to the same four datasets but instead of ranking species by persistence, they were ranked by abundance alone. This was because in many datasets, based on single sampling events, it is not possible to rank according to persistence. Species near boundaries were compared and if of equal abundance the boundary was adjusted accordingly.

2.5 | Application of new sequential Simpson's index method applied to nine published datasets

The new method was applied to nine datasets from seven papers. These were as follows.

1. The average abundance and number of sites occupied by *Onthophagus* (Coleoptera: Scarabaeidae) in Sarawak, Malaysia, grouped as core or satellite by Hanski (1982). Data extracted from figures 4 and 9 in Hanski (1982).
2. Tree species abundance from Barro Colorado Island (BCI), Panama, classified as common or rare by Gray et al. (2005). Species groupings were derived from the BCI-50 plot in figure 3b of Gray et al., 2005, with original species data from Condit et al. (2012).

3. Carabidae (Coleoptera) abundance and number of sites found in Poland. Data from table 1 of Ulrich and Zalewski (2006).
4. The total abundance and number of days sampled for tintinnid ciliates (Ciliophora: Choreotrichia: Tintinnida) in the Mediterranean Sea. Data extracted from figure 3 of Dolan et al. (2009).
5. The persistence and abundance of fish species from an artificial marine reef in California, USA, which were divided into core and transient groups by Boisnier et al. (2010). Original data from table 5 of Matthews (1985).
6. Boisnier et al. (2010) also used a second dataset, from a study of fish from an artificial marine reef in Australia. The data from table 2 in Branden et al. (1994).
7. Two datasets, phytoplankton and fish from a lake in Wisconsin, USA, were divided into core, intermediate, and occasional species by Hansen and Carey (2015). Data from supporting information S1 table in Hansen and Carey (2015).
8. Abundance data of small mammals from Arizona, USA, divided into core, intermediate, and transient by Supp et al. (2015). Data were available from table 1 in Supp et al. (2015).

All were ordered by the frequency of species being found. For data from Hanski (1982), BCI (Gray et al. 2005), and Ulrich & Zalewski (2006) this was in terms of the number of sites a species was present, with the number of times a species was found for the remainder.

3 | RESULTS

3.1 | Effect of group size on classification of core species—comparison of methods

With the Genner et al. (2004) method, the number of core species classified showed a strong positive relationship with the size of the grouping used as a measure of persistence (Spearman rank correlation: Hemiptera $r_s = 13.22$, $n = 10$, $p < 0.001$; Coleoptera $r_s = 2.51$, $n = 10$, $p < 0.001$) (Figure 1). For Hemiptera, the smallest group size (mean 43.6 individuals per group) resulted in six species and the largest (500.9) 18 species, classified as core (Figure 1a). For Coleoptera,

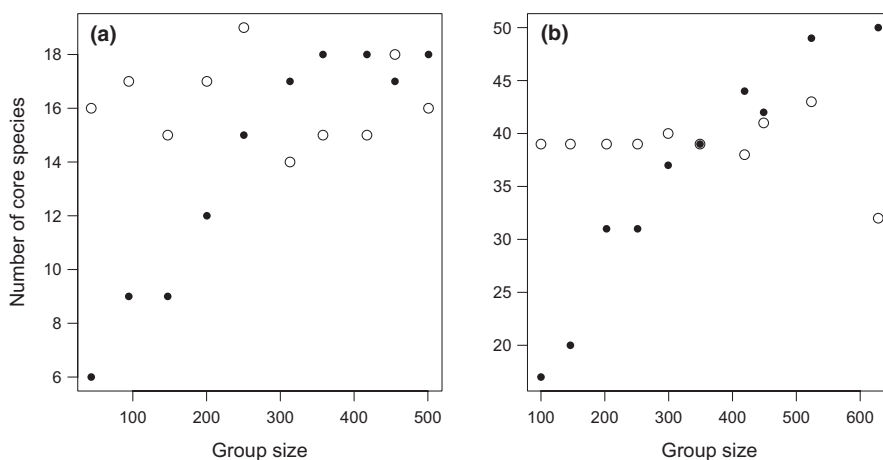


FIGURE 1 The effect of group size on the number of species classified as core species, using the technique of Genner et al. (2004) (filled circles) and the proposed new sequential Simpson's index method (open circles); with group size being the number of individuals allocated to a group to enable the calculation of persistence (frequency of occurrence). Data come from samples from Teagasc Grange: (a) Hemiptera; (b) Coleoptera

the equivalent figures were 17 (100.0) and 50 (628.3) (Figure 1b). In contrast, there was no significant effect of group size using the propose new sequential Simpson's index method (Spearman rank correlation: Hemiptera $r_s = 177.22$, $n = 10$, $p < 0.839$; Coleoptera $r_s = 157.53$, $n = 10$, $p = 0.901$) (Figure 1).

3.2 | Proposed new sequential Simpson's index method applied to different community datasets

The following description of the results focuses on the overall patterns, particularly in terms of the core–occasional species split. More detailed description of the model output for each dataset can be found in Appendix A.2.

3.3 | Bristol Channel fish

When ranked by persistence, the community was divided between 33 core species within four groups, and 48 occasional species in five groups (Table 1, Figure 2a), which was identical to the 33–48 split of Genner et al. (2004). When ranked by abundance, 30 species were classified as core (in three groups) and 51 as occasional (two groups). The ratio of core and occasional species, when ordered by persistence and by abundance using the new method, showed no significant difference ($\chi^2_1 = 0.234$, $p = 0.629$). The abundance-based division was also not significantly different than that of Genner et al. ($\chi^2_1 = 0.234$, $p = 0.629$). Although there was little difference between the core–occasional split between persistence ranked and abundance ranked data, the subgroups did show differences in size and number (Figures 2a, 3a).

3.4 | English Channel fish

Modeling of persistence ranked data resulted in 39 core species in four groups and 33 occasional species also in four groups (Table 1, Figure 2b), which was not significantly different from the 33–39 split of Genner et al. ($\chi^2_1 = 1.000$, $p = 0.317$). The division between

core and occasional species with was not significantly different from data ordered by abundance (41 core and 31 occasional species) ($\chi^2_1 = 0.113$, $p = 0.737$) or from the classification of Genner et al. (2004) ($\chi^2_1 = 1.779$, $p = 0.182$). The number and size of species groupings were very similar between persistence ranked and abundance ranked data (Figures 2b and 3b).

3.5 | Grange Hemiptera

The persistence ranked community was split into six core groups containing 16 species and three occasional groups with a total of 25 species (Table 1).

The core group with the greatest persistence contained a single species of grass-feeding aphid, *Rhopalosiphum* that was clearly more abundant and more frequent than all other species (Appendix A.3). The next most persistent core group also contained a single grass-feeding aphid, *Metopolophium*, much less abundant but still found in 102 of the 115 groups. Further details of the boundaries of each core and occasional group, in terms of persistence and abundance figures, are shown in Appendix A.3.

One species of *Javesella* (Delphacidae) was in the fourth core group, another in the fifth, and a third was in the occasional group with the highest persistence. Similarly, the genus *Macrosteles* (Cicadellidae) had one species in the fourth core, one species in fifth core, and one in sixth core group. Details of the species within each group can be found in Appendix A.4.

Modeling of abundance ranked data resulted in classification of 15 core and 26 occasional species. There was no significant difference between the number of species categorized as either core or occasional between the persistence ranked and abundance ranked data ($\chi^2_1 = 0.203$, $p = 0.652$). There was also no difference in the categorization of the subgroup size (Fisher's exact test $p = 0.889$).

3.6 | Grange Coleoptera

The best fitting model for persistence ranked data divided the community into five species groups: two core groups, with a total of 38

TABLE 1 Species groupings of the two fish and two insect communities deconstructed using the sequential Simpson's index method, after being ranked by persistence

	Species grouping											% of individuals in core groups
	C1	C2	C3	C4	C5	C6	O1	O2	O3	O4	O5	
Bristol Channel fish	11	6	8	8	–	–	7	14	9	5	13	99.4%
English Channel fish	6	10	11	12	–	–	10	11	6	6	–	99.7%
Grange Hemiptera	1	1	2	3	4	5	5	11	9	–	–	98.5%
Grange Coleoptera	21	17	–	–			21	22	44	–		94.7%

Note: C = core species, O = occasional species.

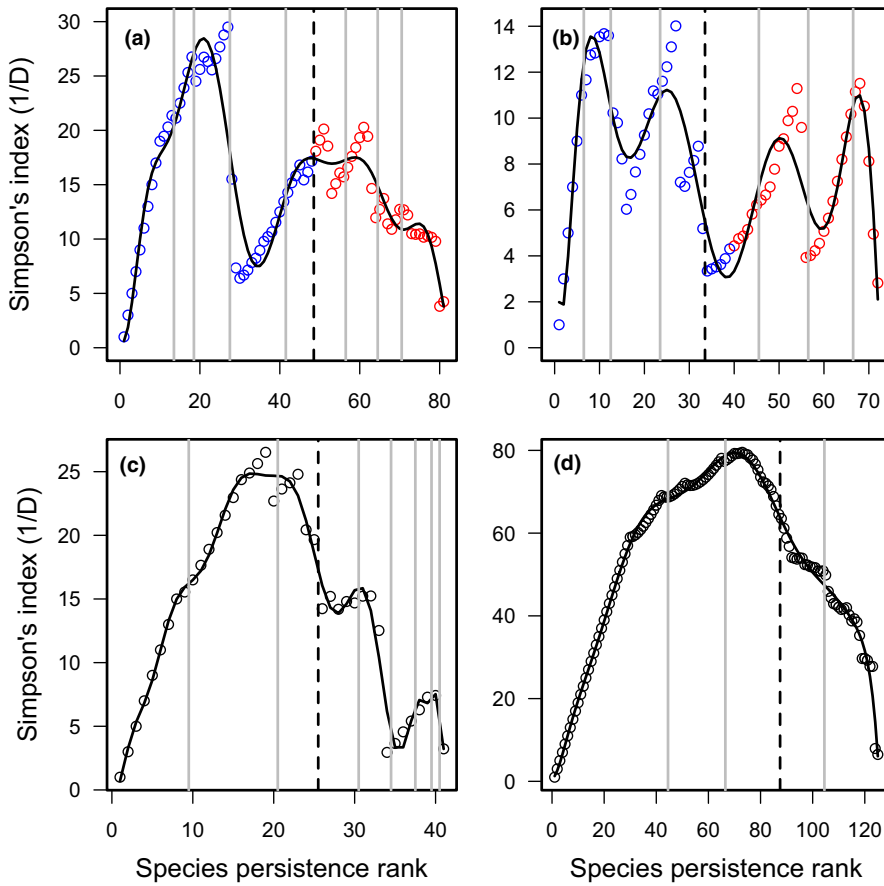


FIGURE 2 Simpson's index ($1/D$) values calculated from species abundance data ranked by persistence (frequency of occurrence) with the sequential removal of the most persistent species. Solid lines show the best fitting polynomial models. Vertical lines show the boundaries between species groups, with the dashed line indicating the division between core and occasional species groups. Species groupings indicated by Genner et al. (2004) are shown as red for core and blue for occasional. (a) Bristol Channel fish; (b) English Channel fish; (c) Grange Hemiptera; (d) Grange Coleoptera

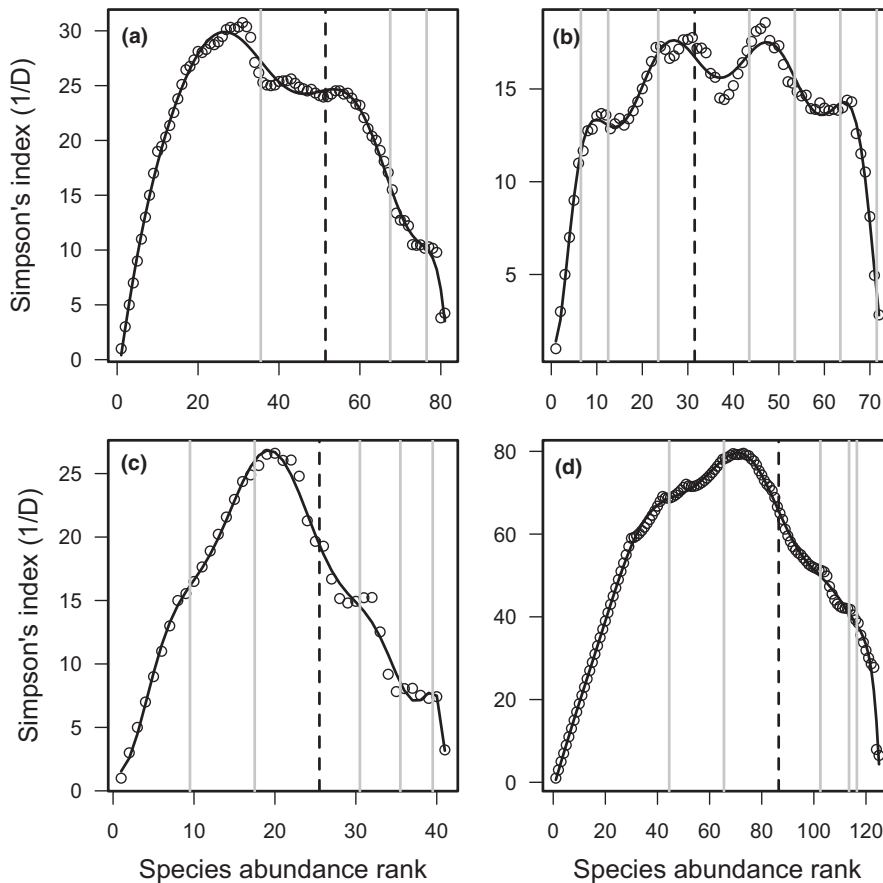


FIGURE 3 Simpson's index ($1/D$) values calculated from species abundance data ranked by abundance with the sequential removal of the most abundant species. Lines show the best fitting polynomial models. (a) Bristol Channel fish; (b) English Channel fish; (c) Grange Hemiptera; (d) Grange Coleoptera

species, and three occasional, with 87 species (Table 1). Details of the boundaries of each core and occasional group, in terms of persistence and abundance figures, are shown in Appendix A.3.

Two species, *Mocyta fungi* and *Amischa analis*, occurred in all possible sample groups and were both far more abundant than any other species. There were 11 species from the genus *Stenus* (Staphylinidae). Four in the most persistent core group, three in the next group, three in the first occasional group, and one in the next (Appendix A.4). There were three Curculionidae that specialize on feeding on *Trifolium*: *Sitona lepidus* and *Protapion fulvipes* in the most persistent core group and *Ischnopterapion virens* in the most frequent occasional group. Details of species groupings can be found in Appendix A.4.

The distribution of species into core or occasional categories was not significantly different from when data were ranked by persistence ($\chi^2_1 = 0.019$, $p = 0.891$). The distribution of subgroups was almost identical for occasional groups and the least abundant/persistent core group but for the abundance ranked data there were a further three core groups as opposed to one for persistence ranked.

3.7 | *Onthophagus* (Coleoptera), Malaysia (Hanski, 1982)

The community was divided into eight core (in three groups) and 10 occasional species (in two groups) of *Onthophagus*, which was not significantly different ($\chi^2_1 = 0.111$, $p = 0.739$) from the nine core and nine satellite species identified by Hanski (1982) (Table 2, Figure 4a).

3.8 | Tree species, Barro Colorado Island (BCI), Panama (Gray et al., 2005)

The 225 species were divided into four core groups, totaling 138 species, and three occasional groups with 92 species (Table 2). If the most occasional group was considered equivalent to the rare category of Gray et al. (2005), which is logical given the rapid rise in Simpson's index value for species rank 1–39 shown in the polynomial model (Figure 4b), and the rest were classified as common, there was no significant difference between the two approaches ($\chi^2_1 = 0.227$, $p = 0.633$).

3.9 | Carabidae (Coleoptera), Poland (Ulrich & Zalewski, 2006)

There were 44 carabid species grouped into six core groups and 31 species in the two occasional groups (Table 2, Figure 4c). Ulrich & Zalewski (2006) suggested 20 core and 31 satellite species, with 24 considered intermediate. If the core groups were considered equivalent to the categories adopted by these authors as core and intermediate species (total 44), then the groupings are identical in size between the two approaches.

3.10 | Tintinnid ciliates, Mediterranean Sea (Dolan et al., 2009)

The community was divided into three core groups, totaling 30 species and two occasional groups, which together had 18 species (Table 2, Figure 4d). Dolan et al. (2009) suggested 11 core and 49 occasional species (the disparity in overall totals being due to undetected data points from figure 3 of Dolan et al. (2009), the source of the data). There was a significant difference in the groupings using the two techniques ($\chi^2_1 = 22.088$, $p < 0.001$). However, the division between Dolan's core and occasional aligns almost exactly with the boundary between groups C1 and C2. Dolan et al. (2009) considered core species to be only those that were found on all 18 days sampled, whereas with my approach species found on 7 or more days were included.

3.11 | Artificial reef fish, California (Boisnier et al., 2010)

The Californian fish community was divided into 16 core species, in four groups, and five occasional species in a single group (Table 2, Figure 3e). This was exactly the same core–occasional classification as used by Boisnier et al. (2010).

3.12 | Artificial reef fish, Australia (Boisnier et al., 2010)

There was a total of 22 species of Australian fish in five core groups, and 25 species divided among the three occasional groups (Table 2, Figure 3f). This was not significantly different from the 29 core and 18 transient species suggested by Boisnier et al. (2010) ($\chi^2_1 = 2.100$, $p = 0.147$) (Figure 4f).

3.13 | Lake fish, Wisconsin, USA (Hansen & Carey, 2015)

Of the 36 species of fish, 22 were core, arranged in four groups, with 15 occasional species in three groups (Table 2, Figure 4g). This was significantly different from 11 core and 25 intermediate and occasional as classified by Hansen and Carey (2015) ($\chi^2_1 = 5.625$, $p = 0.018$) (Table 2, Figure 4g). However, C1–C3 total 13 species (Table 2), which is very close to Hansen & Carey's core classification.

3.14 | Lake phytoplankton, Wisconsin, USA (Hansen & Carey, 2015)

The phytoplankton community was divided into 90 core species, in three groups, and 156 occasional species in two groups (Table 2, Figure 4h). This was very different from the 10 core species suggested by Hansen and Carey (2015). However, the least frequent

TABLE 2 Species groupings of nine additional datasets deconstructed using the sequential Simpson's index method

Original study	Species grouping from original study	Species grouping of present study								
		C1	C2	C3	C4	C5	C6	O1	O2	O3
Hanski (1982)	Core = 9 Sat. = 9	2	3	3	–	–	–	5	5	–
Gray et al. (2005) ^c	Com. = 183 Rare = 42	29	33	40	36	–	–	27	26	39
Ulrich and Zalewski (2006)	Core = 20 Int. = 24 Sat. = 31	2	4	6	10	10	12	13	18	–
Dolan et al. (2009)	Core = 11 Occ. = 49	3	9	18	–	–	–	11	7	–
Boisnier et al. (2010) ^a (USA)	Core = 16 Trans. = 5	1	2	4	9	–	–	5	–	–
Boisnier et al. (2010) ^b (Australia)	Core = 29 Trans. = 18	1	3	5	6	7	–	7	6	12
Hansen and Carey (2015) (fish spp.)	Core = 11 Int. = 23 Occ. = 2	2	6	5	8	–	–	5	6	4
Hansen and Carey (2015) (phytoplankton spp.)	Core = 10 Int. = 146 Occ. = 97	35	21	34	–	–	–	67	96	–
Supp et al. (2015)	Core = 10 Int. = 4 Trans. = 7	1	3	4	–	–	–	4	9	–

Note: C = core species, O = occasional, Com. = common, Int. = intermediate; Occ. = occasional; Sat. = satellite; Trans. = transient.

^aUsing data from Matthews (1985).

^bUsing data from Branden et al. (1994).

^cUsing data from BCI Condit et al. (2012).

occasional grouping was of 96 species, which was almost identical to the 97 considered occasional by the source study.

3.15 | Small mammals, Arizona, USA (Supp et al., 2015)

There were three core subgroups containing a total of eight species and two occasional groups, with 13 species (Table 2, Figure 4i). If the occasional groups were considered to be equivalent to the intermediate and transient species of Supp et al. (2015), then the groupings from the two techniques were almost identical ($\chi^2_1 = 0.097$, $p = 0.755$).

4 | DISCUSSION

Previous plotting of abundance against persistence (i.e., frequency of occurrence) used to deconstruct communities into core and occasional species (Genner et al., 2004; Magurran & Henderson, 2003) relied on the availability of community data over many sampling events, and without this, it has been difficult to reliably distinguish these species groups (Barlow et al., 2010). However, the strong effect of group size on the number of species classified as being core

makes it difficult to know which group size is most appropriate and to have confidence in the allocation of species. Moreover, this problem does not just apply to short-term datasets but to any sample series. For example, if with the dataset from Genner et al. (2004) persistence were measured in months, half years or pairs of years instead of on an annual basis, the result using these previous methods would be a large difference in the number of core species identified. So, does that mean that the classification of core species by Magurran and Henderson (2003) and by Genner et al. (2004) was unreliable? Although that is theoretically possible, in actuality this appeared not to be the case, as the division of the Bristol Channel fish community was found to be closely aligned to the ecology of the species and their relative abundance (Magurran & Henderson, 2003). However, the problem of group size still puts into question the applicability of grouping techniques that rely on persistence in this way.

The group size problem was solved with the new method by using persistence simply as a ranking, not as a continuous variable. The community was then deconstructed based on the differential abundance of species using the property of Simpson's index ($1/D$) as a measure of the probability that any two individuals drawn at random are the same species (Magurran, 2004). Using Simpson's index on progressively smaller sections of the community is a novel approach. Boisnier et al. (2010) had done something similar before but added all species at a given level of persistence, whereas I had removed

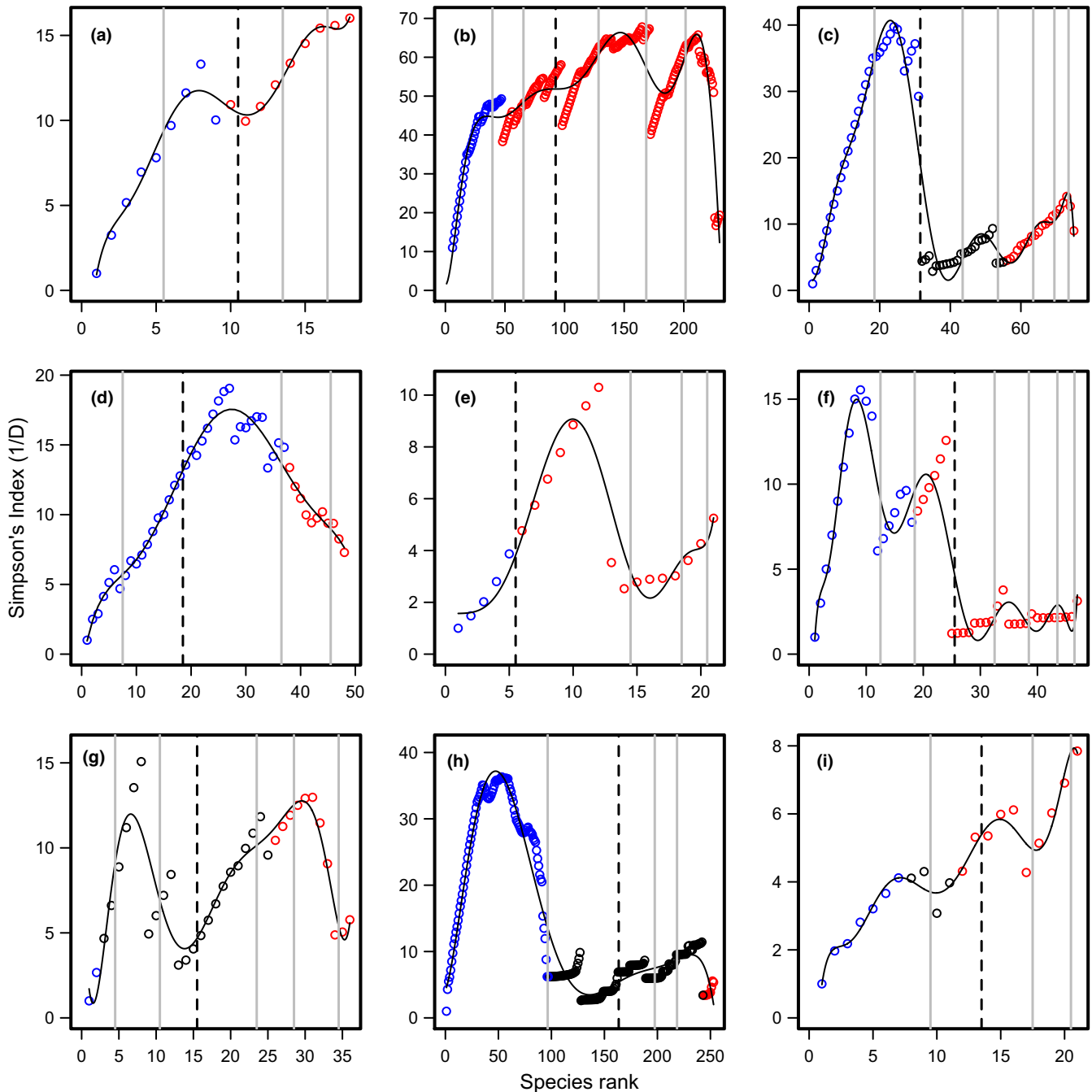


FIGURE 4 Simpson's index ($1/D$) values calculated from species abundance data, from nine previously studied datasets, ranked by persistence (except (i), which was ranked abundance) with the sequential removal of the most persistent (abundant) species. Solid lines show the best fitting polynomial models. Vertical lines show the boundaries between species groups, with the dashed line indicating the division between suggested core and occasional species groups. Species groupings identified by previous studies are shown as red for core, blue for occasional and black for intermediate species. Data from: (a) Hanski (1982); (b) Condit et al. (2012), divided into two groups by Gray et al. (2005); (c) Ulrich and Zalewski (2006); (d) Dolan et al. (2009); (e) Matthews (1985), grouped as core-occasional by Boisnier et al. (2010); (f) Branden et al. (1994), grouped by Boisnier et al. (2010); (g) fish data from Hansen and Carey (2015); (h) phytoplankton data from Hansen and Carey (2015); and (i) Supp et al. (2015)

individual species, irrespective of whether they had the same persistence as the next species or not. They had also been expecting a positive, linear regression, whereas I did not have any a priori expectation of a particular model and was interested in describing the shape of the relationship and what it indicated about species groups.

However, a really important outcome is that, on the assumption, a community is adequately sampled, there is no need for many years of data to divide it into core and occasional species. A particularly useful consequence of which is that by assessing the identity of species groupings over a series of shorter time periods, changes to

community structure over time, including the status of individual species, can be more easily studied (Genner et al., 2004; Henderson & Magurran, 2014; Henderson et al., 2011; Magurran, 2007; Magurran & Henderson, 2010). It could help to show whether the number or proportion of species in a group changed or whether particular species switched between subgroups or from a core group to an occasional grouping or vice versa. Such changes might be related to species invasions, climate change, or other anthropogenic factors such as changing habitat management and the differential removal of species through hunting.

The sequential Simpson's index method resulted in a very similar or identical pattern of deconstruction previously identified for the Bristol Channel (Magurran & Henderson, 2003) and English Channel fish communities (Genner et al., 2004), indicating it is a reliable way of distinguishing core species from occasional. However, more than simply dividing communities into two groups the higher order polynomials of the models, with their multiple inflection points, resulted in core and occasional species, being divided into a number of subgroups. Therefore, the new method enables both a binary division and a more complex structure that may be a better reflection of reality.

One of the main advantages of the new method is that it reduces subjectivity relative to other methods of grouping core and occasional species, through identifying an optimum model based on statistical grounds. Previous methods, such as those of Hansen and Carey (2015) and Dolan et al. (2009), have had a more subjective approach, in their cases considering core species only as those that were found on all sampling occasions. Application of the new method is much more objective and is likely to be much closer to a representation of reality.

One of the criticisms of the concept of a core (i.e., common)-occasional (i.e., satellite or rare) dichotomy has been that it is an oversimplification of reality (Ulrich & Ollik, 2004; Whittaker, 2015). The new method retains the core-occasional framework while identifying of subgroups helps to describe better the patterns of species that occur between the extremes of persistence (Ugland & Gray, 1982; Ulrich & Ollik, 2004). This approach may be particularly useful in studying multimodal species abundance distributions (Antão et al., 2017; Matthews, Borges, et al., 2014).

Do subgroups have any biological or ecological meaning or are they simply a description of probability due to relative abundance of their component species? Magurran and Henderson (2003) found that in an estuarine ecosystem the core species were those associated with muddy substrates, whereas the occasional species preferred habitats, or are normally found in deep water. This showed that there was a connection between habitat association, abundance, and persistence. So, by extension, habitat preference or other ecological characteristics can be expected to align with subgroups. Indeed, Ugland and Gray (1982) suggested that communities are composed of groups of species whose constituents have some similarity in their adaptation to a habitat. Therefore, subgroups of species may have ecological meaning in relation to habitat preference and other niche characteristics.

Unfortunately for many species, particularly invertebrates, knowledge of their ecology is limited, so it may be difficult to relate this to species groupings. However, some comment can be made about the subgroups of Hemiptera and Coleoptera identified from the grassland at Grange. The Hemiptera genera, *Javesella* and *Macrosteles*, were represented by relatively unspecialized grass-feeding species, often associated with disturbed grasslands (Nickel, 2003). Within the subgroups, both genera showed a clear abundance ranking of species. A similar pattern was found in Coleoptera of the genus *Stenus*, which are predators of soft-bodied arthropods such as Collembola, and in *Trifolium* feeding weevils (Curculionioidea) (Lott & Anderson, 2011; Morris, 1990, 2002). The order and scale of the differences between these species may reflect differences their respective niches (Harpole & Tilman, 2006; Southwood, 1996). Many other core Hemiptera species were generalist grass feeders, and core Coleoptera were often generalist predators or associated with cattle dung (Nickel, 2003; Skidmore, 1991). Several core species could be related to the presence of their specialized food plants, such as the aphid *Thecabius affinis*, which feeds of *Ranunculus*, and the aphid *Acyrtosiphon pisum* a specialist on *Trifolium repens* and other leguminous plants (Blackman, 2010; Heie, 1980). Consequently, these Hemiptera and Coleoptera examples suggest that the subgrouping of species may well be related to their ecological niche.

The identity of the species included in the subgroups did differ depending on whether persistence or abundance was used to rank them. However, due to the close relationship between persistence and abundance (Magurran, 2004; Magurran & Henderson, 2003), the overall size and pattern of the subgroups showed little difference between models based on abundance ranking and those with persistence ranking. Consequently, abundance data alone could be used for the identification of core and occasional species and of the subgroup structure. To do so, a community would have to be well sampled and data give a good representation of true species richness (Gotelli & Colwell, 2001; Magurran, 2004). However, on that assumption, the new technique does not require long runs of sample data over many years. Therefore, data collected over short-term periods or which are related to very long-lived species such as trees (Barlow et al., 2010; Condit et al., 2012; Umaña et al., 2017) can be deconstructed. Furthermore, communities could be deconstructed using data from individual years allowing investigation of how species groupings change over time in relation to issues such as climate change and human impact (Genner et al., 2004; Henderson & Magurran, 2014; Henderson et al., 2011; Magurran, 2007; Magurran & Henderson, 2010). Allocation of a species to a grouping would be mathematically independent of their status in other years. This would be particularly useful in studying changes in increasing or declining species, such as the fish *Liparis liparis* in the Bristol Channel, which may have declined due to increasing water temperature (Henderson et al., 2011). The new approach could also be extended to spatial studies to identify core and satellite species and species subgroupings, assessed at the individual site scale (Matthews, Borges, et al., 2014; Supp et al., 2015).

The application of the new sequential Simpson's method to nine further datasets investigated its wider applicability. The division into core and occasional was identical or very nearly identical to the classification used by Hanski (1982), Boisnier et al. (2010) (Matthews 1985 data), Ulrich & Zalewski (2006) and Supp et al. (2015). For that of Gray et al. (2005), Dolan et al. (2009), and Boisnier et al. (2010) (Brandon 1994 data), binary core–occasional division did differ but there were strong similarities in other aspects of the groups identified. The least similarity was shown with the two Hansen and Carey (2015) datasets, who had not made a binary division but instead had grouped species at the extremes of abundance and/or persistence as core and occasional, with those species between being considered intermediate. Therefore, the new method can be successfully applied to a range of taxa, locations, and study approaches.

The division of communities into multiple groupings is aligned with the view that a binary division into core and occasional is simplistic and in reality communities are much more likely to be multimodal in structure (Antão et al., 2017; Matthews, Borges, et al., 2014). Although the new method can be used for a simple core–occasional dichotomy, it also identifies a more complex structure of multimodality. By applying the multiple groups to the previously studied datasets, rather than focusing on a simple dichotomy, reveals that there is strong agreement between the findings of the new approach and the various techniques used by other studies, even when the initial core–occasional grouping did not fit as well.

5 | CONCLUSION

In applying the core–occasional species concept to a short-term data set of grassland insects, I found a probability based way of dividing communities into species groups. I have demonstrated that by sequentially removing species and calculating Simpson's index reveals patterns that distinguish the same core and occasional species classification as proposed by earlier studies and so show that it can be applied to a wide range of datasets, whether based on short- or long-term data. However, it has the advantage over previous methods in having no a priori assumption of a binary division of species but rather can identify multimodality in species abundance distributions. It agrees well with previous studies that have used a range of taxa and sampling methodologies and offers a more objective approach to the study of species groupings. Moreover, it does not rely on long time series of sample data and can be used even if only species abundances are known, assuming well-sampled communities. It can therefore be applied to a far greater range of community data allowing a more fine-scaled approach and so has the potential to provide a valuable insight into how communities change in time or space.

ACKNOWLEDGMENTS

I am very grateful to Peter Henderson and Anne Magurran for giving me permission to use their fish community data and to Condit et al. (2012) for the use of BCI data. Annette Anderson helped with collection and initial sorting into orders of some of the insect samples.

I thank Michael Drennan and the Teagasc Animal & Grassland Research Centre, Grange, for permitting access to their suckler beef grassland experiments. Fabrizio Manco gave initial help with some of the R code used. Insect sampling was part of the Ag-Biota Project, funded by the Environmental Protection Agency, Ireland (2001-CD/B1-M1) through the ERTDI Programme under the National Development Plan (2000–2006), Ireland. The Ag-Biota Project was planned and led by Gordon Purvis. The subsequent work did not receive any grants from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The author has no competing interests to declare.

AUTHOR CONTRIBUTION

Alvin J. Helden: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Project administration (lead); Resources (lead); Software (lead); Validation (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (lead).

DATA AVAILABILITY STATEMENT

Data from Teagasc Grange can be accessed at <https://doi.org/10.5061/dryad.ksn02v74g>.

Data from the Bristol Channel fish study are available from the Pisces Conservation website at: <http://www.pisces-conservation.com/phmm-data-old.html>.

The Bristol Channel and English Channel data are available online in an appendix to Genner et al. (2004). Permission was also given by Peter Henderson and Anne Magurran to use their fish data.

Sources of data used from other published studies are described in the Section 2.

ORCID

Alvin J. Helden  <https://orcid.org/0000-0002-8607-8356>

REFERENCES

- Antão, L. H., Connolly, S. R., Magurran, A. E., Soares, A., & Dornelas, M. (2017). Prevalence of multimodal species abundance distributions is linked to spatial and taxonomic breadth: Multimodal SADs linked to spatial and taxonomic breadth. *Global Ecology and Biogeography*, 26(2), 203–215. <https://doi.org/10.1111/geb.12532>
- Astudillo-García, C., Bell, J. J., Webster, N. S., Glasl, B., Jompa, J., Montoya, J. M., & Taylor, M. W. (2017). Evaluating the core microbiota in complex communities: A systematic investigation: Core microbiota in complex communities. *Environmental Microbiology*, 19(4), 1450–1462. <https://doi.org/10.1111/1462-2920.13647>
- Barbosa, A. M., Real, R., Munoz, A. R., & Brown, J. A. (2015). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, 19(10), 1333–1338. <https://onlinelibrary.wiley.com/doi/full/10.1111/ddi.12100>
- Barlow, J., Gardner, T. A., Louzada, J., & Peres, C. A. (2010). Measuring the conservation value of tropical primary forests: The Effect of occasional species on estimates of biodiversity uniqueness. *PLoS One*, 5(3), e9609. <https://doi.org/10.1371/journal.pone.0009609>
- Barnes, C. J., Burns, C. A., van der Gast, C. J., McNamara, N. P., & Bending, G. D. (2016). Spatio-temporal variation of core and satellite arbuscular

- mycorrhizal fungus communities in *Miscanthus giganteus*. *Frontiers in Microbiology*, 7, 1278. <https://doi.org/10.3389/fmicb.2016.01278>
- Blackman, R. L. (Ed.) (2010). *Handbooks for the identification of British insects. Vol. 2, Pt. 7: Aphids – Aphidinae (Macrosiphini)*. Field Studies Council.
- Boisnier, E., Sagawa, T., Komatsu, T., & Takagi, N. (2010). Distinguishing resident from transient species along marine artificial reefs. *Journal of the Marine Biological Association of the United Kingdom*, 90(7), 1297–1303. <https://doi.org/10.1017/S0025315410000512>
- Boss, R. L., & Silva, J. M. C. (2014). Core and transient species in an Amazonian savanna bird assemblage. *Revista Brasileira De Ornitologia*, 22, 374–382. <https://doi.org/10.1007/BF03544274>
- Branden, K. L., Pollard, D. A., & Reimers, H. A. (1994). A review of recent artificial reef developments in Australia. *Bulletin of Marine Science*, 55, 982–994.
- Condit, R., Lao, S., Pérez, R., Dolins, S. B., Foster, R. B., & Hubbell, S. P. (2012). *Barro Colorado Forest Census Plot Data, 2012 Version*. Center for Tropical Forest Science Databases. <https://doi.org/10.5479/data.bci.20130603>
- Coyle, J. R., Hurlbert, A. H., & White, E. P. (2013). Opposing mechanisms drive richness patterns of core and transient bird species. *The American Naturalist*, 181(4), E83–E90. <https://doi.org/10.1086/669903>
- Crawley, M. J. (2007). *The R book*. John Wiley & Sons Ltd.
- Dolan, J. R., Ritchie, M. E., Tunin-Ley, A., & Pizay, M.-D. (2009). Dynamics of core and occasional species in the marine plankton: Tintinnid ciliates in the north-west Mediterranean Sea. *Journal of Biogeography*, 36(5), 887–895. <https://doi.org/10.1111/j.1365-2699.2008.02046.x>
- Dornelas, M., & Connolly, S. R. (2008). Multiple modes in a coral species abundance distribution. *Ecology Letters*, 11(10), 1008–1016. <https://doi.org/10.1111/j.1461-0248.2008.01208.x>
- Drennan, M. J., & McGee, M. (2009). Performance of spring-calving beef suckler cows and their progeny to slaughter on intensive and extensive grassland management systems. *Livestock Science*, 120(1–2), 1–12. <https://doi.org/10.1016/j.livsci.2008.04.013>
- Genner, M. J., Sims, D. W., Wearmouth, V. J., Southall, E. J., Southward, A. J., Henderson, P. A., & Hawkins, S. J. (2004). Regional climatic warming drives long-term community changes of British marine fish. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271(1539), 655–661.
- Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379–391. <https://doi.org/10.1046/j.1461-0248.2001.00230.x>
- Gray, J. S., Bjorgesaeter, A., & Ugland, K. I. (2005). The impact of rare species on natural assemblages. *Journal of Animal Ecology*, 74(6), 1131–1139. <https://doi.org/10.1111/j.1365-2656.2005.01011.x>
- Hansen, G. J. A., & Carey, C. C. (2015). Fish and phytoplankton exhibit contrasting temporal species abundance patterns in a dynamic north temperate lake. *PLoS One*, 10(2), e0115414. <https://doi.org/10.1371/journal.pone.0115414>
- Hanski, I. (1982). Dynamics of regional distribution—The core and satellite species hypothesis. *Oikos*, 38(2), 210–221. <https://doi.org/10.2307/3544021>
- Harpole, W. S., & Tilman, D. (2006). Non-neutral patterns of species abundance in grassland communities. *Ecology Letters*, 9(1), 15–23. <https://doi.org/10.1111/j.1461-0248.2005.00836.x>
- Heie, O. E. (1980). *The Aphidoidea (Hemiptera) of Fennoscandia and Denmark. I. General Part. The families Mindaridae, Hormaphididae, Thelaxidae, Anoeciidae, and Pemphigidae*. Fauna Entomologica Scandinavica (Vol. 9, 236pp.). Scandinavian Science Press Ltd.
- Helden, A. J., Anderson, A., Finn, J., & Purvis, G. (2015). The response of sward-dwelling arthropod communities to reduced grassland management intensity in pastures. *Irish Journal of Agricultural and Food Research*, 54(2), 107–120. <https://doi.org/10.1515/ijagr-2015-0012>
- Henderson, P. A., & Magurran, A. E. (2014). Direct evidence that density-dependent regulation underpins the temporal stability of abundant species in a diverse animal community. *Proceedings of the Royal Society B: Biological Sciences*, 281(1791), 20141336. <https://doi.org/10.1098/rspb.2014.1336>
- Henderson, P. A., Seaby, R. M. H., & Somes, J. R. (2011). Community level response to climate change: The long-term study of the fish and crustacean community of the Bristol Channel. *Journal of Experimental Marine Biology and Ecology*, 400(1–2), 78–89. <https://doi.org/10.1016/j.jembe.2011.02.028>
- Lande, R., DeVries, P. J., & Walla, T. R. (2000). When species accumulation curves intersect: Implications for ranking diversity using small samples. *Oikos*, 89, 601–605. <https://doi.org/10.1034/j.1600-0706.2000.890320.x>
- Lott, D. A., & Anderson, R. (2011). *The Staphylinidae (rove beetles) of Britain and Ireland. Parts 7 & 8: Oxyporinae, Steninae, Euaesthetinae, Pseudopsinae, Paederinae, Staphylininae*. Handbooks for the identification of British insects (Vol. 12 part 7, pp. 1–340). The Royal Entomological Society and the Field Studies Council.
- Magurran, A. E. (2004). *Measuring biological diversity*. Blackwell Publishing.
- Magurran, A. E. (2007). Species abundance distributions over time. *Ecology Letters*, 10(5), 347–354. <https://doi.org/10.1111/j.1461-0248.2007.01024.x>
- Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422, 714–716. <https://doi.org/10.1038/nature01547>
- Magurran, A. E., & Henderson, P. A. (2010). Temporal turnover and the maintenance of diversity in ecological assemblages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1558), 3611–3620. <https://doi.org/10.1098/rstb.2010.0285>
- Matthews, K. R. (1985). Species similarity and movement of fishes on natural and artificial reefs in Monterey Bay, California. *Bulletin of Marine Science*, 37, 252–270.
- Matthews, T. J., Borges, P. A. V., & Whittaker, R. J. (2014). Multimodal species abundance distributions: A deconstruction approach reveals the processes behind the pattern. *Oikos*, 123(5), 533–544. <https://doi.org/10.1111/j.1600-0706.2013.00829.x>
- Matthews, T. J., Borregaard, M. K., Ugland, K. I., Borges, P. A. V., Rigal, F., Cardoso, P., & Whittaker, R. J. (2014). The gambin model provides a superior fit to species abundance distributions with a single free parameter: Evidence, implementation and interpretation. *Ecography*, 37(10), 1002–1011. <https://doi.org/10.1111/ecog.00861>
- Morris, M. G. (1990). *Orthocerous Weevils. Coleoptera Curculionoidea (Nemonychidae, Anthribidae, Urodontidae, Attelabidae and Apionidae)*. Handbooks for the identification of British insects (Vol. 5 part 16, pp. 1–108). Royal Entomological Society of London.
- Morris, M. G. (2002). *True Weevils (Part 1). Coleoptera: Curculionidae (Subfamilies Raymondionyminae to Smicronychinae)*. Handbooks for the identification of British insects (Vol. 15 part 17(b), pp. 1–149). Royal Entomological Society of London and the Field Studies Council.
- Nickel, H. (2003). *The leafhoppers and planthoppers of Germany (Hemiptera, Auchenorrhyncha): Patterns and strategies in a highly diverse group of phytophagous insects*. Pensoft.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Rodwell, J. S. (2006). *National vegetation classification: Users' handbook*. Joint Nature Conservation Committee.
- Skidmore, P. (1991). *Insects of the British Cow Dung Community*. Field Studies Council.
- Southwood, T. R. E. (1996). The Croonian lecture, 1995—Natural communities: Structure and dynamics. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 351(1344), 1113–1129.
- Supp, S. R., Koons, D. N., & Ernest, S. K. M. (2015). Using life history trade-offs to understand core-transient structuring of a small mammal community. *Ecosphere*, 6(10), art187. <https://doi.org/10.1890/ES15-00239.1>

- Ugland, K. I., & Gray, J. S. (1982). Lognormal distributions and the concept of community equilibrium. *Oikos*, 39(2), 171–178. <https://doi.org/10.2307/3544482>
- Ugland, K. I., Lamshead, P. J. D., McGill, B., Gray, J. S., O'Dea, N., Ladle, R. J., & Whittaker, R. J. (2007). Modelling dimensionality in species abundance distributions: Description and evaluation of the Gambin model. *Evolutionary Ecology Research*, 9, 313–324.
- Ulrich, W., & Ollik, M. (2004). Frequent and occasional species and the shape of relative-abundance distributions. *Diversity and Distributions*, 10(4), 263–269. <https://doi.org/10.1111/j.1366-9516.2004.00082.x>
- Ulrich, W., & Zalewski, M. (2006). Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. *Oikos*, 114(2), 338–348. <https://doi.org/10.1111/j.2006.0030-1299.14773.x>
- Umaña, M. N., Zhang, C., Cao, M., Lin, L., & Swenson, N. G. (2017). A core-transient framework for trait-based community ecology: An example from a tropical tree seedling community. *Ecology Letters*, 20(5), 619–628. <https://doi.org/10.1111/ele.12760>
- van der Gast, C. J., Walker, A. W., Stressmann, F. A., Rogers, G. B., Scott, P., Daniels, T. W., Carroll, M. P., Parkhill, J., & Bruce, K. D. (2011). Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *The ISME Journal*, 5(5), 780–791. <https://doi.org/10.1038/ismej.2010.175>
- Whittaker, R. J. (2015). On the species abundance distribution in applied ecology and biodiversity management. *Journal of Applied Ecology*, 52(2), 443–454. <https://doi.org/10.1111/1365-2664.12380>
- Winterbottom, J. M. (1949). Mixed bird parties in the tropics, with special reference to Northern Rhodesia. *The Auk*, 66, 258–263. <https://doi.org/10.2307/4080356>

How to cite this article: Helden, A. J. (2021). Core and occasional species: A new way forward. *Ecology and Evolution*, 11, 10547–10565. <https://doi.org/10.1002/ece3.7863>

APPENDIX A

Appendix A1: R code used to apply the sequential Simpson's index method, followed by method used to identify model inflection points

Part 1: R code

The following are data of species abundances of fish from the Bristol Channel. The data come from Magurran and Henderson (2003) and are ordered by the number of years found in the community (persistence)*. The data are placed in vector a. [If there is no persistence/frequency ranking known, or if only relative abundance is of interest, the values need to be ranked i.e. `a<-sort(a)`]

```
a<-c(1,1,1,1,1,1,1,1,1,1,2,2,2,3,2,2,2,5,3,3,5,6,4,4,4,5,24,57,75,5,9,15,7,15,18,7,7,19,22,40,17,19,11,36,88,17,31,25,34,40,109,198,49,61,172,53,8-8,124,60,64,220,444,601,127,238,865,943,303,366,1012,1342,2189,2236,2642,3544,3926,5122,6815,32692,37379)
```

One is added to each value in vector a. This is to avoid multiplying by zero when all remaining species in a sub-set of the community only contains singletons.

```
a=a+1
```

The following code creates a series of sequentially smaller vectors, with the most abundant species being removed at each step. So, the first vector contains all species, the second contains all species except the most abundant, the third all species except the two most abundant etc. The final vector created has only the final two species in the original dataset (in most cases these would be singletons). The vectors are placed a matrix called listOutput.

```
b=length(a)-1
listOutput<-vector("list", length(b))
for (i in 0:b)
{
  if (mean(a[1:(length(a)-i)]))
  {
    output<-a[1:(length(a)-i)]
    listOutput[[i+1]]<-output
  }
}
```

This code firstly creates a function to calculate Simpson's index (1/D). It then calculates the Simpson's index value for each species vector and places the output into a new vector called simp.vect2. Lastly it creates a vector called spp, which is simply a list of integers from 1 to the number of values of Simpson's index calculated.

```
SIndex<-function(x) {
  1/(sum((x*(x-1))/((sum(x))*(sum(x)-1))))
}
simp.vect2<-sapply(listOutput,SIndex)
spp<-c(length(a):1)
```

The Simpson's index values (simp.vect2) can then be modelled with spp as the explanatory variable, for example as follows.

```
model9x<-glm(simp.vect2~spp+l(spp^2)+l(spp^3)+l(spp^4)+l(spp^5)+l(spp^6)+l(spp^7)+l(spp^8)+l(spp^9),family=gaussian)
summary(model9x)
```

The data can be plotted with the following code to show a scatter plot with the linear model added.

```
plot(spp,simp.vect2,las=1,xlab="Species persistence rank",ylab="Simpson's index (1/D)",pch=16)
points(spp,(((coef(model9x)["l(spp^9)"]*(spp^9))+((coef(model9x)["l(spp^8)"]*(spp^8))+((coef(model9x)["l(spp^7)"]*(spp^7))+((coef(model9x)["l(spp^6)"]*(spp^6))+((coef(model9x)["l(spp^5)"]*(spp^5))+((coef(model9x)["l(spp^4)"]*(spp^4))+((coef(model9x)["l(spp^3)"]*(spp^3))+((coef(model9x)["l(spp^2)"]*(spp^2))+((coef(model9x)["spp"])*spp+(coef(model9x)["(Intercept)"]))),type="l",las=1)
```

To enable identification of inflection points in the model, the equation for the model can be solved for y for each species rank (x)

```
y<-(((coef(model9x)["l(spp^9)"]*(spp^9))+((coef(model9x)["l(spp^8)"]*(spp^8))+((coef(model9x)["l(spp^7)"]*(spp^7))+((coef(model9x)["l(spp^6)"]*(spp^6))+((coef(model9x)["l(spp^5)"]*(spp^5))+((coef(model9x)["l(spp^4)"]*(spp^4))+((coef(model9x)["l(spp^3)"]*(spp^3))+((coef(model9x)["l(spp^2)"]*(spp^2))+((coef(model9x)["spp"])*spp+(coef(model9x)["(Intercept)"]))))
```

[*The values for persistence (number of years recorded) by which the abundance data in vector a were ordered are: 1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,3,3,3,3,4,4,4,4,4,4,4,5,5,5,6,6,6,7,7,7,7,7,8,8,9,9,9,10,10,13,13,14,15,16,17,17,18,19,19,19,20,20,20,20,20,21,21,21,21,22,22,22,22,22,22,22,22,22,22,22,22]

Part 2: Identification of inflection points

With higher order polynomials, the inflection point cannot be determined by differentiation, therefore the model data need to be investigated as follows.

- The values of the model (y) for each species rank (x) are listed (see table, below – data from Supp et al. (2015)).
- The difference is taken between each value of y and the value of the preceding value i.e. $y_x - y_{x-1}$.
- The difference values change sequentially in either a positive or negative direction.
- Inflection points are represented when the difference values change direction. In the table below, the superscript letters (a-g) indicate this, with inflection points occurring between values with different letters. For example, the difference figure decreases from species 2 to species 3, then after species 3 it starts to increase (until species 5), so an inflection point (change from a decreasing difference to increasing difference) lies somewhere between species 3 and 4.

Species rank (x)	Value of model (y)	Difference between y_x and y_{x-1}
1	0.979	
2	2.040	1.061 ^a (i.e., 2.040–0.979)
3	2.184	0.144 ^a
4	2.610	0.426 ^b
5	3.292	0.683 ^b
6	3.876	0.584 ^c
7	4.118	0.242 ^c
8	4.022	–0.096 ^c
9	3.787	–0.235 ^c
10	3.675	–0.112 ^d
11	3.871	0.196 ^d
12	4.393	0.522 ^d
13	5.076	0.683 ^d
14	5.643	0.567 ^e
15	5.838	0.195 ^e
16	5.588	–0.250 ^e
17	5.120	–0.468 ^e
18	4.953	–0.167 ^f
19	5.649	0.696 ^f
20	7.184	1.536 ^f
21	7.790	0.605 ^g

Appendix A2: Details of final models for the 13 datasets used in the study

Bristol Channel fish

When ranked by persistence a 14th order polynomial showed the best fit. The glm model was as follows:

$$y = 3.593e^{-1} + 3.345e^{-1}x^3 - 9.269e^{-2}x^4 + 1.160e^{-2}x^5 - 8.424e^{-4}x^6 + 3.912e^{-5}x^7 - 1.220e^{-6}x^8 + 2.620e^{-8}x^9 - 3.883e^{-10}x^{10} + 3.909e^{-12}x^{11} - 2.553e^{-14}x^{12} + 9.764e^{-17}x^{13} - 1.661e^{-19}x^{14}$$

In this model, there were 10 inflection points, but the three least persistent were combined giving a division of the community into nine groups. Four groups were considered more persistent, core groups, together containing 99.4% of all individuals, and there were five less persistent, occasional groups.

When ranked by abundance, data were best described by a ninth order polynomial. The glm model was as follows:

$$y = -2.111 + 2.531x - 1.313e^{-2}x^3 + 1.239e^{-3}x^4 - 5.846e^{-5}x^5 + 1.523e^{-6}x^6 - 2.206e^{-8}x^7 + 1.663e^{-10}x^8 - 5.081e^{-13}x^9$$

Deconstructing the community resulted in three core groups (C1 = 5, C2 = 9, C3 = 16) totaling 30 species and 99.4% of individuals, and two occasional groups (O1 = 16, O2 = 35) totaling 51 species.

English Channel fish

The best fitting model for persistence ranked data was a tenth order polynomial. The community was split into four core groups, containing 99.7% of all individuals, and four occasional groups. The glm model was as follows:

$$y = 6.488 - 7.674x + 3.746x^2 - 0.628e^{-1}x^3 + 5.263e^{-2}x^4 - 2.552e^{-3}x^5 + 7.564e^{-5}x^6 - 1.391e^{-6}x^7 + 1.549e^{-8}x^8 - 9.556e^{-11}x^9 + 2.508e^{-13}x^{10}$$

The abundance rank data were best described with a tenth order polynomial model. The glm model was as follows:

$$y = 2.413 - 2.620x + 1.912x^2 - 3.391e^{-1}x^3 + 2.949e^{-2}x^4 - 1.470e^{-3}x^5 + 4.474e^{-5}x^6 - 8.451e^{-7}x^7 + 9.671e^{-9}x^8 - 6.139e^{-11}x^9 + 1.659e^{-13}x^{10}$$

The deconstructed community grouping was for five core groups (C1 = 1, C2 = 8, C3 = 10, C4 = 10, C5 = 12), totaling 41 species and 99.9% of individuals and four occasional groups (O1 = 8, O2 = 11, O3 = 6, O4 = 6), with 31 species.

Grange Hemiptera

The model that best fitted the persistence ranked data was a 16th order polynomial. The glm model was as follows:

$$y = -3.125e^{-1} + 2.238x^3 - 1.736x^4 + 6.260e^{-1}x^5 - 1.336e^{-1}x^6 + 1.861e^{-2}x^7 - 1.787e^{-3}x^8 + 1.218e^{-4}x^9 - 5.989e^{-6}x^{10} + 2.131e^{-7}x^{11} - 5.436e^{-9}x^{12} + 9.688e^{-11}x^{13} - 1.145e^{-12}x^{14} + 8.066e^{-15}x^{15} - 2.562e^{-17}x^{16}$$

The community was split into six core groups containing 98.5% of individuals and three occasional groups.

The abundance ranked data were best described with a tenth order polynomial. The glm model was as follows:

$$y = 1.336 + 2.714e^{-1}x^3 + 7.350e^{-2}x^4 + 8.777e^{-3}x^5 - 5.773e^{-4}x^6 + 2.230e^{-5}x^7 - 5.043e^{-7}x^8 + 6.186e^{-9}x^9 - 3.181e^{-11}x^{10}$$

The community was deconstructed into four core groups (C1 = 2, C2 = 5, C3 = 4, C4 = 4) totaling 15 species and representing 98.2% of the individuals, and three occasional groups (O1 = 10, O2 = 10, O3 = 6), with 26 species overall.

Grange Coleoptera

The best fitting model for persistence ranked data was an 11th order polynomial. The glm model was as follows:

$$y = 7.362e^{-1} + 6.010e^{-1}x^2 - 7.751e^{-2}x^3 + 5.274e^{-3}x^4 - 2.091e^{-4}x^5 + 5.097e^{-6}x^6 - 7.866e^{-8}x^7 + 7.711e^{-10}x^8 - 4.655e^{-12}x^9 + 1.579e^{-14}x^{10} - 2.306e^{-17}x^{11}$$

There were seven inflection points, and the community was divided into five species groups: two core groups and three occasional. The two core groups contained 94.7% of individuals.

A 12th order model best described the pattern in the abundance ranked data. The glm model was as follows:

$$y = -1.550 + 2.142x - 2.508e^{-4}x^4 + 3.125e^{-5}x^5 - 1.662e^{-6}x^6 + 4.796e^{-8}x^7 - 8.237e^{-10}x^8 + 8.684e^{-12}x^9 - 5.523e^{-14}x^{10} + 1.948e^{-16}x^{11} - 2.931e^{-19}x^{12}$$

There were four core groups (C1 = 9, C2 = 3, C3 = 11 and C4 = 16) totaling 39 species and including 95.0% of all individuals, and three occasional groups (O1 = 21, O2 = 21, O3 = 44).

Onthophagus (Coleoptera), Malaysia (Hanski, 1982)

$$y = -7.846 + 14.480x - 7.533x^2 + 2.098x^3 - 3.031e^{-1}x^4 + 2.308e^{-2}x^5 - 8.799e^{-4}x^6 + 1.325e^{-5}x^7$$

Tree species, Barro Colorado Island (BCI), Panama (Gray et al., 2005)

$$y = 1.390 + 3.362e^{-1}x^2 - 2.110e^{-2}x^3 + 6.029e^{-4}x^4 - 9.743e^{-6}x^5 + 9.570e^{-8}x^6 - 5.816e^{-10}x^7 + 2.134e^{-12}x^8 - 4.323e^{-15}x^9 + 3.710e^{-18}x^{10}$$

Carabidae (Coleoptera), Poland (Ulrich & Zalewski, 2006)

$$y = -1.328 + 2.060e^{-1}x^3 - 4.792e^{-2}x^4 + 5.040e^{-3}x^5 - 3.007e^{-4}x^6 + 1.107e^{-5}x^7 - 2.608e^{-7}x^8 + 3.942e^{-9}x^9 - 3.703e^{-14}x^{10} + 1.970e^{-13}x^{11} - 4.535e^{-12}x^{12}$$

Tinitinnid ciliates, Mediterranean Sea (Dolan et al., 2009)

$$y = -1.058 + 2.340x - 3.631e^{-1}x^2 + 3.046e^{-2}x^3 - 1.157e^{-3}x^4 + 1.982e^{-5}x^5 - 1.262e^{-7}x^6$$

Artificial reef fish, California (Boisnier et al., 2010)

$$y = 1.569 + 1.900e^{-3}x^5 - 4.183e^{-4}x^6 + 3.439e^{-5}x^7 - 1.251e^{-6}x^8 + 1.700e^{-8}x^9$$

Artificial reef fish, Australia (Boisnier et al., 2010)

$$y = -10.690 + 20.220x - 11.650x^2 + 3.492x^3 - 5.659e^{-1}x^4 + 5.405e^{-2}x^5 - 3.225e^{-3}x^6 + 1.237e^{-4}x^7 - 3.051e^{-6}x^8 + 4.678e^{-8}x^9 - 4.060e^{-10}x^{10} + 1.524e^{-12}x^{11}$$

Lake fish, Wisconsin, USA (Hansen & Carey, 2015)

$$y = 9.970 - 13.930x + 7.086x^2 - 1.356x^3 + 1.298e^{-1}x^4 - 6.905e^{-3}x^5 + 2.078e^{-4}x^6 - 3.315e^{-6}x^7 + 2.180e^{-8}x^8$$

Lake phytoplankton, Wisconsin, USA (Hansen & Carey, 2015)

$$y = 4.063 + 4.585e^{-1}x + 4.782e^{-2}x^2 - 1.555e^{-3}x^3 + 1.791e^{-5}x^4 - 9.939e^{-8}x^5 + 2.701e^{-10}x^6 - 2.893e^{-13}x^7$$

Small mammals, Arizona, USA (Supp et al., 2015)

$$y = -4.953 + 10.810x - 6.615x^2 + 2.051x^3 - 3.449e^{-1}x^4 + 3.290e^{-2}x^5 - 1.776e^{-3}x^6 + 5.049e^{-5}x^7 - 5.866e^{-7}x^8$$

Appendix A3: Persistence range, measured as the number of times species were present in samples (maximum value shown in column title), also expressed as a percentage (in brackets), and abundance range shown by each of the species groups identified by deconstructing the Hemiptera and Coleoptera communities from Grange

Species grouping	Hemiptera		Coleoptera	
	Persistence/115 (%)	Abundance	Persistence/63 (%)	Abundance
C1	115 (100)	2,672	29–63 (46.0–100)	43–1,957
C2	102 (88.7)	551	15–28 (23.8–44.4)	16–33
C3	71–80 (61.7–69.6)	168–224	–	–
C4	54–66 (47.0–57.4)	117–345	–	–
C5	21–53 (18.3–46.1)	26–456	–	–
C6	11–20 (9.6–17.4)	24–28	–	–
O1	3–10 (2.6–8.7)	7–11	5–14 (7.9–22.2)	5–17
O2	2–3 (1.7–2.6)	2–7	3–4 (4.7–6.3)	3–6
O3	1 (0.9)	1–2	1–2 (1.6–3.2)	1–3

Appendix A4: List of species of Hemiptera and Coleoptera from Grange, contained within each group identified by deconstructing the communities. Within each group, species are ranked first by persistence (frequency), starting with the most persistent at the top, and then alphabetically

Hemiptera species	Hemiptera group	Coleoptera species	Coleoptera group
<i>Rhopalosiphum</i> Koch sp.	C1	<i>Mocyta fungi</i> (Gravenhorst)	C1
<i>Metopolophium</i> Mordvilko sp.	C2	<i>Amischa analis</i> (Gravenhorst)	C1
<i>Myzus</i> Passerini sp.B	C3	<i>Aloconota gregaria</i> (Erichson)	C1
<i>Myzus</i> Passerini sp.A	C3	<i>Stenus formecitorum</i> Mannerheim	C1
<i>Javesella obscurella</i> (Boheman)	C4	<i>Stenus nanus</i> Stephens	C1
<i>Thecabius affinis</i> (Kaltenbach)	C4	<i>Sitona lepidus</i> (Fabricius)	C1
<i>Macrosteles viridigriseus</i> (Edwards)	C4	<i>Ptenidium pusillum</i> (Gyllenhal)	C1
<i>Sipha glyceriae</i> (Kaltenbach)	C5	<i>Megasternum concinnum</i> (Marsham)	C1
<i>Javesella pellucida</i> (Fabricius)	C5	<i>Tachyporus pusillus</i> Gravenhorst	C1
<i>Sitobion avenae</i> (Fabricius)	C5	<i>Stenus clavicornis</i> (Scopoli)	C1
<i>Macrosteles sexnotatus</i> (Fallén)	C5	<i>Acrotichis atomaria</i> (De Geer)	C1
<i>Microlophium evansi</i> (Theobald)	C6	<i>Amischa decipiens</i> (Sharp)	C1
<i>Sitobion fragariae</i> (Walker)	C6	<i>Pterostichus strenuus</i> Panzer	C1
<i>Acyrtosiphon pisum</i> (Harris)	C6	<i>Atomaria nitidula</i> (Marsham)	C1
<i>Macrosteles laevis</i> (Ribaut)	C6	<i>Stenus picipes</i> Stephens	C1
<i>Uromelan</i> (Mordvilko) sp.	C6	<i>Bembidion lampros</i> (Herbst)	C1
<i>Saldula orthochila</i> (Fieber)	O1	<i>Tachyporus chrysomelinus</i> (L.)	C1
<i>Atheroides serrulatus</i> Haliday	O1	<i>Amischa nigrofusca</i> (Stephens)	C1
<i>Javesella dubia</i> (Kirschbaum)	O1	<i>Ptenidium nitidum</i> (Heer)	C1
<i>Anthocoris nemorum</i> (L.)	O1	<i>Longitarsus luridus</i> (Scopoli)	C1

Hemiptera species	Hemiptera group	Coleoptera species	Coleoptera group
<i>Aphrodes albifons</i> (L.)	O1	<i>Protapion fulvipes</i> (Geoffroy)	C1
<i>Holcaphis</i> Hille Ris Lambers sp.	O2	<i>Philonthus carbonarius</i> (Gravenhorst)	C2
Unidentified aphid sp.G	O2	<i>Stenus cicindeloides</i> (Schaller)	C2
<i>Philaenus spumarius</i> (L.)	O2	<i>Acrotoma</i> Thomson, C.G. sp.A	C2
<i>Pithanus maerkeli</i> (Herrich-Schäffer)	O2	<i>Stenus brunripes</i> Stephens	C2
<i>Phyllaphis fagi</i> (L.)	O2	<i>Bembidion guttula</i> (Fabricius)	C2
<i>Nabis ferus</i> (L.)	O2	<i>Oxyptoda</i> Mannerheim sp.E	C2
<i>Myzus cerasi</i> (Fabricius)	O2	<i>Gabrieus appendiculatus</i> Sharp	C2
<i>Drymus sylvaticus</i> (Fabricius)	O2	<i>Stenus similis</i> (Herbst)	C2
<i>Brachycaudus helichrysi</i> (Kaltenbach)	O2	<i>Tachyporus dispar</i> (Paykull)	C2
<i>Aphrodes makarovi</i> Zakhvatkin	O2	<i>Pterostichus vernalis</i> (Panzer)	C2
<i>Aphis</i> L. sp.	O2	<i>Ptiliolum spencei</i> (Allibert)	C2
<i>Pachytomella parallela</i> (Meyer-Dür)	O3	<i>Tachyporus hypnorum</i> (Fabricius)	C2
<i>Streptanus sordidus</i> (Zetterstedt)	O3	<i>Acrotichis grandicollis</i> (Mannerheim)	C2
<i>Stenodema calcarata</i> (Fallén)	O3	<i>Bembidion aeneum</i> Germar	C2
<i>Saldula saltatoria</i> (L.)	O3	<i>Philonthus cognatus</i> Stephens	C2
<i>Psylla melanoneura</i> Förster	O3	<i>Tachyporus nitidulus</i> (Fabricius)	C2
<i>Paraliburnia clypealis</i> (J. Sahlberg)	O3	<i>Atomaria apicalis</i> Erichson	C2
<i>Megophthalmus scabripennis</i> Fallén	O3	<i>Stenus canaliculatus</i> Gyllenhal	O1
<i>Aphalara exilis</i> (Weber & Mohr)	O3	<i>Cartodere nodifer</i> (Westwood)	O1
<i>Acyrtosiphon</i> Mordvilko sp.A	O3	<i>Stenus ossium</i> Stephens	O1
		<i>Helophorus brevipalpis</i> Bedel	O1
		<i>Atomaria atricapilla</i> Stephens	O1
		<i>Stephostethus lardarius</i> (De Geer)	O1
		<i>Quedius schatzmayri</i> Gridelli	O1
		<i>Ceutorhynchus erysimi</i> (Fabricius)	O1
		<i>Anotylus tetracaratus</i> (Block)	O1
		<i>Oligota</i> Mannerheim sp.	O1
		<i>Stenus fulvicornis</i> Stephens	O1
		<i>Ischnoptera virens</i> (Herbst)	O1
		<i>Nebria brevicollis</i> (Fabricius)	O1
		<i>Cercyon melanocephalus</i> L.	O1
		<i>Xantholinus linearis</i> (Olivier)	O1
		<i>Gabrieus breviventer</i> (Sperk)	O1
		<i>Enicmus histrio</i> Joy & Tomlin	O1
		<i>Xantholinus longiventris</i> Heer	O1
		<i>Tachinus laticollis</i> Gravenhorst	O1
		<i>Philonthus laminatus</i> (Creutzer)	O1
		<i>Cercyon impressus</i> (Sturm)	O1
		<i>Agonum muelleri</i> (Herbst)	O2
		<i>Trechus quadristriatus</i> (Schrank)	O2
		<i>Tachyporus obtusus</i> (L.)	O2
		<i>Rhinoncus pericarpus</i> (L.)	O2
		<i>Pteryx suturalis</i> (Heer)	O2
		<i>Philonthus marginatus</i> (Müller)	O2
		<i>Loricera pilicornis</i> (Fabricius)	O2

Hemiptera species	Hemiptera group	Coleoptera species	Coleoptera group
		<i>Dimetrota</i> Mulsant & Rey sp.B	O2
		<i>Dimetrota</i> Mulsant & Rey sp.A	O2
		<i>Datomicra</i> Mulsant & Rey sp.A	O2
		<i>Clivina fossor</i> (L.)	O2
		<i>Cercyon pygmaeus</i> (Illiger)	O2
		<i>Tachinus rufipes</i> (L.)	O2
		<i>Leiosoma deflexum</i> (Panzer)	O2
		<i>Coccidula rufa</i> (Herbst)	O2
		<i>Stenus junco</i> (Paykull)	O2
		<i>Philhygra</i> Mulsant & Rey sp.C	O2
		<i>Microdota</i> Mulsant & Rey sp.B	O2
		<i>Corticara gibbosa</i> (Herbst)	O2
		<i>Ceutorhynchus typhae</i> (Paykull)	O2
		<i>Autalia rivularis</i> (Gravenhorst)	O2
		<i>Aleochara lanuginosa</i> Gravenhorst	O2
		<i>Platystethus arenarius</i> (Fourcroy)	O3
		<i>Malthodes pumilus</i> (Brébisson)	O3
		Unknown Aleocharinae sp.A	O3
		<i>Tachyporus solutus</i> Erichson	O3
		<i>Pterostichus melanarius</i> (Illiger)	O3
		<i>Ptenidium fuscicorne</i> Erichson	O3
		<i>Propylea quattuordecimpunctata</i> (L.)	O3
		<i>Oxytelus laqueatus</i> (Marsham)	O3
		<i>Microdota</i> Mulsant & Rey sp.E	O3
		<i>Lema cyanella</i> (L.)	O3
		<i>Encephalus complicans</i> Stephens	O3
		<i>Cypha</i> Leach sp.	O3
		<i>Cercyon haemorrhoidalis</i> (Fabricius)	O3
		<i>Anotylus rugosus</i> (Fabricius)	O3
		<i>Xantholinus punctulatus</i> (Paykull)	O3
		<i>Typhaea stercorea</i> (L.)	O3
		<i>Tychus niger</i> (Paykull)	O3
		<i>Tachyporus tersus</i> Erichson	O3
		<i>Rugilus similis</i> (Erichson)	O3
		<i>Quedius nitipennis</i> (Stephens)	O3
		<i>Philonthus varians</i> (Paykull)	O3
		<i>Philonthus intermedius</i> (Lacordaire)	O3
		<i>Philhygra</i> Mulsant & Rey sp.D	O3
		<i>Oxypoda</i> Mannerheim sp.C	O3
		<i>Oxypoda</i> Mannerheim sp.A	O3
		<i>Omonadus floralis</i> (L.)	O3
		<i>Ischnosoma splendidum</i> (Gravenhorst)	O3
		<i>Micropeplus porcatus</i> (Paykull)	O3
		<i>Micrambe vini</i> (Panzer)	O3
		<i>Meligethes aeneus</i> (Fabricius)	O3
		<i>Megarthus depressus</i> (Paykull)	O3

Hemiptera species	Hemiptera group	Coleoptera species	Coleoptera group
		<i>Megarthus denticollis</i> (Beck)	O3
		<i>Datomicra</i> Mulsant & Rey sp.D	O3
		<i>Chaetida longicornis</i> (Gravenhorst)	O3
		<i>Cercyon nigriceps</i> (Marsham)	O3
		<i>Carpelimus bilineatus</i> (Stephens)	O3
		<i>Cantharis nigra</i> (De Geer)	O3
		<i>Calathus melanocephalus</i> (L.)	O3
		<i>Brachypterus urticae</i> (Fabricius)	O3
		<i>Barypeithes pellucidus</i> (Boheman)	O3
		<i>Atomaria umbrina</i> (Gyllenhal)	O3
		<i>Atheta</i> Thomson, C.G. sp.B	O3
		<i>Amara aenea</i> (De Geer)	O3
		<i>Acrotrichis sericans</i> (Motschulsky)	O3