Full title: Multivariate methods for testing hypotheses of temporal community dynamics

Running header: Testing hypotheses of temporal community dynamics

Authors and affiliations:

Hannah L. Buckley^{1*}, Nicola J. Day², Bradley S. Case¹, Gavin Lear³, Aaron M. Ellison⁴

¹ School of Science, Auckland University of Technology, Auckland, New Zealand

² Biology Department, Wilfrid Laurier University, Waterloo, Ontario, Canada

³ School of Biological Sciences, University of Auckland, Auckland, New Zealand

⁴ Harvard Forest, Harvard University, Petersham, Massachusetts, USA

* Corresponding author: Hannah.Buckley@aut.ac.nz

ABSTRACT

For ecological research to make important contributions towards understanding and managing temporally-variable global change processes, such as responses to land-use and climatic change, we must have effective and comparable ways to quantify and analyse compositional change over time in biological communities. These changes are the sum of local colonisation and extinction events, or changes in the biomass and relative abundance of taxa within and among samples. We conducted a quantitative review of currently available methods for the analysis of multivariate datasets collected at temporal intervals. This review identified the need for the application of quantitative, hypothesis-based approaches to understand temporal change in community composition, particularly for small datasets with less than 15 temporal replicates. To address this gap, we: (1) conceptually present how temporal patterns in community dynamics can be framed as specific, testable hypotheses; (2) provide three fully-worked case-studies, complete with R code, demonstrating multivariate analysis methods for temporal hypothesis testing and pattern visualisation; and (3) present a road map for testing specific, quantitative hypotheses relating to the underlying mechanisms of temporal community dynamics.

KEYWORDS

Beta diversity, biodiversity, community variation, compositional change, multivariate analysis, species turnover, temporal change, temporal variability, time series, zeta diversity

INTRODUCTION

A key task for ecologists is to monitor, quantify, analyse, and predict temporal changes in the structure of ecological communities as a function of ongoing changes in climate, land use, and other environmental drivers. Ecological communities comprise populations of multiple taxa. The structure of communities—the number of taxa (‘richness’) and their relative abundance (‘composition’)—changes through time because of fluctuations of individual

populations and local colonisations and extinctions of taxa (Micheli et al. 1999, Legendre and Gauthier 2014). These temporal dynamics are governed by both intrinsic factors, including intraspecific and interspecific interaction, colonisation, and extinction, and extrinsic factors, which are processes that relate to periodic disturbances and directionally changing environmental conditions (Connell and Slatyer 1977, Brown and Lawson 2010). Understanding the temporal dynamics of community structure can illuminate fundamental ecological processes, including effects of individual life histories on rates of ecosystem change, the relative importance of biotic and abiotic factors in determining community structure, or how taxa and the networks in which they are embedded respond to environmental change (Tilman 1999, Hobbs et al. 2007, Kampichler et al. 2012). Temporal changes in taxon richness and composition also can provide both simple summaries of ecosystem change and information that can be translated rapidly into action, either to protect and restore declining taxa, or to staunch unsustainable growth of unwanted organisms.

Representations of community structure as multivariate datasets that include information on multiple taxa can be more sensitive to changes in explanatory or predictor variables (e.g., pH, temperature, soil type), than univariate summary statistics, such as taxon richness, diversity, or biomass (Micheli et al. 1999, Chao et al. 2014), and hence more informative about causal processes. However, quantitative methods for investigating and testing hypotheses about temporal dynamics of multivariate measures of community structure (henceforth ‘community dynamics’) also are more complex than methods used for univariate data (Schaefer et al. 2005, Legendre and Legendre 2012). To date, methods for the analysis of spatial multivariate community structure data (e.g., Whittaker 1960, Legendre and Legendre 2012) have received far more attention than methods for analysis of temporal multivariate data. Well-known concepts used to describe spatial changes in community structure—grain, extent, nestedness, beta diversity, taxonomic turnover—have been

transferred to the temporal domain (Azeria and Kolasa 2008, Anderson et al. 2011, Legendre and Gauthier 2014), and some well-developed multivariate methods for large spatial datasets have been used to detect temporal changes in community structure (Legendre and Gauthier 2014). However, although temporal studies of compositional change are commonly undertaken, it is not always clear which analytical methods are appropriate or available for any given research question, data type, or community structure; indeed, there may not be one ideal method for all situations and a range of methods may be required to answer a given question (Whittaker 1960). Comparisons of multivariate ecological data through time pose unique, non-trivial challenges because these analyses must also account for spatial variation in these data, (Philippi et al. 1998, Micheli et al. 1999, Collins et al. 2000, Schaefer et al. 2005).

In this paper, we: (1) review empirical studies that have undertaken analyses of multivariate temporal community datasets, to gain insight into how researchers have been applying analysis methods over the past several decades, (2) provide recommendations for the greater use of quantitative methods for testing specific temporal hypotheses, and (3) provide guidance for researchers, including those new to community dynamics research, in analysis decisions, particularly for small temporal datasets (fewer than 15 temporal replicates).

How to use this paper

This paper does not necessarily need to be read from start to finish; we suggest readers identify the section(s) most appropriate for their interests. The remainder of this paper is structured into five main sections: (A) a review of the recent literature; (B) a conceptual description of the types of hypotheses that can be addressed in analyses of temporal community dynamics data; (C) three worked case study examples illustrating hypothesis-driven analyses of temporal community data, including code for the software programme ‘R’

(R Core Development Team 2017) to allow readers to easily implement methods described in this paper; (D) A “road map” to analysing temporal community data, including descriptions, and how to identify and present, analyses appropriate for testing particular hypotheses; (E) A brief discussion of recent advances in multivariate analysis of community dynamics.

A. LITERATURE REVIEW

DATA COLLECTION AND ANALYSIS

We reviewed and described the multivariate analyses of community dynamics in studies published up until September 2017 to provide an overview of available methods, their benefits and drawbacks, and what they can reveal about community dynamics. Specifically, we asked: Are data collected from different taxonomic groups, habitats, spatial extents, and time-scales analysed in particular (or different) ways, i.e., are there consistencies in analysis approaches due to discipline-specific preferences?

We searched the literature for studies investigating changes in ecological communities over time. We searched the ISI Web of Science Core Collection on 5 September 2017 for articles using the following search statement: [“temporal” AND “composition*” AND “communit*” AND (“community dynamic*” OR “temporal dynamic*” OR “temporal community varia*” OR “community change” OR “compositional change”)]. We refined this search to include only document types listed as ‘articles’ and published in ‘English’. This search yielded 692 articles for review. Of these, we excluded 241 papers that lacked community data (i.e. those that included fewer than three taxa), had no temporal replication or pooled all temporal data, used only artificial communities or simulated datasets, or used space-for-time substitution (e.g., using chronosequences) to infer community ages. Although we did not constrain the temporal search window, our search terms only returned papers after 1990, indicating that the specific terms we used were not part of the compositional change literature prior to the 1990s.

We extracted and collated key attributes relating to each study in the remaining 451 papers from the ISI database (e.g., publication year, number of citations). Specific study attributes were determined from the body text of each manuscript, including information on (i) habitat type (e.g., estuarine); (ii) habitat location (e.g., U.S.A.); (iii) taxonomic identity (e.g., vertebrate; fish); (iv) manipulation (e.g., natural variation, experimental resource manipulation); (v) spatial extent (e.g., local, global); (vi) temporal grain and extent (e.g., minutes, months); (vii) how composition was measured (e.g., presence-absence, relative abundances of taxa); (viii) how time was included in each analysis (i.e., explicitly as a predictor variable, or where other variables represent time, e.g., climatic factors recorded at different times); and (ix) the key perceived research question (e.g., exploration of yearly dynamics) (Data S1). We categorised organisms being studied into the following taxonomic groups: plants (Kingdom Plantae); vertebrates (Phylum Chordata); invertebrates (Phyla Annelida, Arthropoda, Cnidaria, Ctenophora, Echinodermata, Mollusca, Platyhelminthes, or Porifera); microeukaryotes (algae and Kingdoms Fungi and Protista); prokaryotes (Domains Bacteria and Archaea); and viruses (non-cellular pathogenic organisms). The data types collated and included papers are provided in the supplementary material (Data S1).

LITERATURE REVIEW RESULTS AND DISCUSSION

The 451 papers identified by our search, when compared by year to the number of papers in community ecology, showed that every year since 1990, publications on temporal community dynamics consistently have comprised approximately 24% of studies within community ecology. Thus, studies on temporal dynamics have increased from less than five publications per year in the early 1990s to more than 30 per year in the last five years (Appendix S1: Fig. S1). Few of the 451 reviewed publications (20 papers; 4% of those reviewed) contained more than one independent dataset; those instances were treated as multiple studies that were analysed as separate data points (increasing our total sample size n to 475). The majority of

studies (70%) were conducted within North America and Europe (Fig. 1), likely reflecting the relative availability of research funding and resources.

Reviewed studies generally focused on organisms, scales, and habitats that were conducive to studying temporal community dynamics. For instance, prokaryotes and invertebrates were studied most frequently (Fig. 1), being relatively dynamic and easy to study over smaller temporal and spatial scales; by contrast, research on plants and vertebrate fauna together accounted for less than one-third of all studies of community dynamics. Although studies spanned a broad range of habitat types, taxa, spatial, and temporal scales, there were two-to-three times more studies at small spatial extents (micro, point sample, and local scales) across the 14 habitat categories identified from the articles (Appendix S1: Fig. S2A). Similarly, the most common temporal extent (i.e., the time between the collection of the first and last samples) of individual studies was years, and only a minority of studies on plant and vertebrate communities extended for a decade or more (Appendix S1: Fig. S2B). Smaller organisms were often measured at finer temporal resolutions (grain) than larger organisms, and many studies on short-lived prokaryote and viral communities lasted more than one year. The temporal grain of sampling generally was smaller for communities of microorganisms than for communities of vertebrate animals and plants, for which the most common grain for repeated sample collection was years rather than months or seasons (Appendix S1: Fig. S2B).

Researchers frequently used a range of different analytical methods, even within one study, to enable them to understand the complexities of observed temporal changes. Up to eight different analytical methods were used (Fig. 2A), with about two-thirds of the studies applying ≤ 3 methods, and most using two. Although the number of studies addressing community dynamics have increased dramatically since the 1990s and have begun to incorporate more computationally-complex methods over time (Appendix S1: Fig. S3A),

descriptive methods (i.e., visual, non-statistical comparisons) and ordination have remained the most commonly-used analyses, comprising 33% and 22% of all applied methods, respectively (Fig. 2B). In fact, 134 out of 451 studies (30%) used only descriptive approaches to “analyse” their datasets, while another 35% of studies used descriptive methods in addition to at least one other method. We suspect that the overall ease of use of descriptive and ordination methods, their accessibility via readily-available software programs, and their focus on visually-presenting complex multivariate data, together have driven this trend. The use of the raw dissimilarity values (18%), time-lag analysis (9%), and cluster analysis (7%) comprised the next most common types of methods; the remaining 12 methods were used infrequently (Fig. 2B). Several methods captured by this review were not multivariate (see Appendix 2), including univariate linear methods, temporal stability (*a.k.a.* coefficient of variation), frequency change (testing for significant changes in the frequency of individual taxa over time), and multiplicative change (percent cover change statistic for a single taxon or taxon group used in subsequent regression analyses). Since 2005, the use of raw dissimilarity methods has increased considerably. In the same decade, a steady stream of alternative methods that incorporate time explicitly have been applied, albeit with poor uptake across the wider literature in this area. These methods often are ‘one-off’ analyses that have been applied only a handful of times; we refer to these as ‘minor’ methods (Appendix S1: Fig. S3). Preferred ordination methods changed from detrended correspondence analysis (DCA) and principal components analysis (PCA) in the early 1990s, to canonical correspondence analysis (CCA) in the late 1990s, and, since the early 2000s, to non-metric multidimensional scaling (NMDS) (Appendix S1: Fig. S3B).

Three-quarters of all analyses compiled were from observational studies aimed at describing patterns of community change or related to monitoring community composition, rather than exploring environmental effects or studying the effects of biotic processes or

interactions (Appendix S1: Fig. S4). In fewer cases, researchers were interested in the effect of environmental change, such as seasonality or climatic cycles over longer time periods.

Our review also shows how common it is for researchers to analyse datasets with low temporal replication. Across research aims and for most methods, most time series had fewer than 15 observations (mean = 14.9, median = 7, mode = 4; Fig. 3); the few studies that used network analyses and machine-learning methods used 50 or more observations taken through time. Overall, neither the number of temporal replicates, nor any other study attributes we extracted during our review (e.g., habitat, taxonomic group), appeared to explain well the differences in how researchers analysed their data or their motivations for doing so.

Typically, researchers will have some notion of how their communities of interest might be changing, which motivated them to undertake temporal sampling in the first place (Bestelmeyer et al. 2011). In some cases, long-term datasets exist because previous researchers began data collection with a rationale and method that is different from the current research questions, and so we can sometimes be limited by past sampling designs. Nonetheless, our review highlights that most studies describe temporal community dynamics within their studied communities without explicit, quantitative hypotheses and specific, testable predictions about the expected drivers of observed community changes (Appendix S1: Fig. S4). Clearly, there are many benefits to entering the analysis pathway with a strong *a priori* prediction of the compositional pattern(s) expected and the biological hypotheses that underpin those predictions.

B. HYPOTHESES OF TEMPORAL COMMUNITY DYNAMICS

Quantitative and specific predictions of the nature and variability in temporal community dynamics allows deeper understanding of the patterns and processes involved. However, to achieve this, *a priori* expectations of change need to be developed and paired with the appropriate analytical method(s) to achieve rigorous tests. Here, we describe in detail the

possible observed changes in community composition (taxon presence or abundance) in one sampling unit (site, host, etc.) over time (Fig. 4) or in a set of sampling units over time (Figs. 5-7) and what some of the ecological research questions are that may underpin such hypotheses. In each case, we give predictions for how pairwise dissimilarity values, the basis for most of the commonly-used quantitative measures of compositional change identified in the literature review, would change. Many of these hypotheses are not mutually exclusive although some could be used as multiple alternatives to be tested simultaneously, depending on the research question(s) of interest. Note that ‘dissimilarity measures’, also called ‘distance measures’ or ‘beta-diversity’, are the inverse of ‘similarity measures’. These metrics quantify the compositional distance between two samples (pairwise dissimilarity).

Simple predictions for a single sample unit: Persistence and directional change

Aside from the ‘null’ prediction of no change in composition (‘persistence’; Fig. 4b), one of the simplest predictions of compositional dynamics is that of ‘directional’ change, whereby certain taxa either increase or decrease so that the dissimilarity of a sampling unit to itself at previous sampling occasions increases (or decreases) over time (Fig. 4c); when composition turns over as taxa are replaced, the magnitude of change in dissimilarity can be high. Such a hypothesis could arise from many ecological processes, e.g., has the community undergone ‘succession’, whereby there is a predictable replacement of individual or groups of taxa over time; has it undergone ‘degradation’, whereby rare taxa are lost from the system over time; has it had an ‘invasion’ event, where the relative dominance of one taxon or functional group increases; is it changing in composition in response to a changing biotic or abiotic variable, such as predation or temperature?

Simple patterns for multiple sample units: Directional change, divergence and convergence

Where a set of sample units are all undergoing the same directional changes in composition over time, pairwise ‘spatial’ dissimilarity among that set of sampling units, would not change

significantly over time, because all sampling units are moving in the same direction at the same rate in compositional space because the same taxa are changing in the same ways (Fig. 5b). We also might ask whether sample units are becoming more dissimilar to each other, due to interactions between the changing factors described above and characteristics of the community, such as initial composition or differences in habitat conditions. In this case ‘divergence’ represents the hypothesis where a set of sampling units are all becoming more dissimilar in their composition, so that their pairwise dissimilarities increase over time; this term therefore does not apply to change within one sampling unit (Fig. 5c).

Two very similar patterns predict decreasing pairwise dissimilarity values over time. In contrast to ‘divergence’, under ‘convergence’ taxonomic composition of two or more sampling units becomes more similar over time (Fig. 5e) in a predictable way because of increases or decreases in taxon presence or abundance; this is also called ‘homogenisation’ (Fig. 5e), where there is a concurrent decrease in diversity and dissimilarity becomes relatively low. In contrast, convergence simply implies that between sample-unit dissimilarity declines over time, but not necessarily to a very low level. In the cases where taxa are lost, the predicted overall decline in dissimilarity is large. These hypotheses of change would be relevant if we are interested in the effects of an experimental treatment designed to increase the growth of a particular set of taxa, or under invasion, such as the biotic homogenisation of cities due to urbanisation, or loss of native taxa due to competition with the invading taxa. Note that the predicted effect of invasion on dissimilarity values depends on whether we are considering the trajectory of one sample unit (divergence) or many sample units (convergence) and the among-sample unit variation in other interacting factors.

More complex temporal patterns: Perturbations, resilience, and cyclical change

More complex patterns in temporal dynamics involve compositional shifts that change in their magnitude or direction over time. Natural or manipulative experiments asking questions

about the effects of disturbance on community composition can lead to complex predictions of change. ‘Perturbation’ is a pattern for which the change in composition is initially high but lessens over time and can be detected by comparison to undisturbed treatments or areas (Fig. 6a). The perturbation can occur as either a ‘press perturbation’ (Fig. 6b), such as the removal of a predator or a continuous nutrient addition treatment, or a ‘pulse perturbation’ (Fig. 6c), such as a disturbance event or a population density reduction; these two perturbations result in different rates of change in pairwise temporal dissimilarities. Post-perturbation, a community either could remain in a state that is dissimilar to un-perturbed sample units (‘tipping point’), or could return to its previous state over time (‘resilience’). The prediction of a ‘pivot point’ or ‘threshold’ can be used to test for a significant compositional shift for one or more sampling units over a relatively short period of time, meaning that temporal pairwise dissimilarity among times for each sampling unit is low before, and also after, the pivot point, but is high for temporal pairs across the pivot point (Fig. 6c).

Questions about the scale of temporal dynamics in a community, such as the presence of seasonal cycles, lead to predictions of ‘cyclical’ change or ‘periodicity’, which represent shifts in composition that repeatedly diverge and re-converge back to a previous state; dissimilarity increases at short temporal scales but decreases at larger temporal scales, and will alternate between high and low at increasing temporal scales that correspond to some cyclical process, e.g., warming events (Fig. 4d).

Predictability of change: Instability and stabilisation

Questions about the predictability of compositional changes over time resulting from interactions between causal-change processes, such as temperature shifts throughout a year, and biotic or abiotic differences among sampling units, could also lead to hypotheses regarding the variability in relative compositional change. For instance, ‘instability’ predicts significantly increasing temporal variation in dissimilarity for a set of sampling units because

relative increases and decreases among taxa are not correlated or consistent; change is not directional although mean dissimilarity would also likely increase with the variance (Fig. 5d). In contrast, ‘stabilisation’ predicts the opposite: significantly decreasing temporal variation in dissimilarity as initially significant compositional change lessens over time; pairwise dissimilarity is initially high at short temporal scales (the opposite of Fig. 5d).

Taxon dynamics: Synchronicity

In many cases, we are interested in asking if different components of the community are responding in similar or different ways over time. Where individual taxa, or particular groups of taxa (e.g., functional groups), change in similar ways at the same time (‘synchronicity’), we can detect this because among-taxon-group dissimilarity increases, but within-taxon-group dissimilarity does not change (this involves comparing the pairwise sampling unit dissimilarities separately for different taxonomic subsets) (Fig. 7a). Synchronicity may result in directional or other types of compositional dynamics patterns within one sampling unit or a set of sampling units (Fig. 7b).

Hypotheses of biological processes causing dynamics may be very specific, such as a decrease in certain prey taxa in response to a predator, or more general, such as a shift towards a grassland community in response to fire disturbance of a woodland or an increase in annual plant taxa in spring compared to winter. More specific and directional (increases or decreases) hypotheses lead to more powerful tests of compositional change; we strongly encourage researchers to generate such hypotheses *a priori* and to pair these with predictions that are as quantitative as possible, e.g., define the predicted timescales of variation.

C. WORKED CASE STUDY EXAMPLES

Here, we provide three worked examples to demonstrate the hypothesis-testing approach we advocate for different types of community dynamics. We used open-access datasets to test for: (1) significant divergence in temporal community dynamics among disturbance

treatments imposed upon plant communities in experimental grasslands, (2) neutral community dynamics for a fully-censused forest plot, and (3) seasonal (cyclical) temporal dynamics in marine aquatic bacterial communities. We present a brief description of each dataset and describe the results and interpretation for the relevant exploratory data analysis (EDA; see section ‘D. Road map for analysing temporal community dynamics’) and hypothesis-testing analyses. More detailed descriptions, explanations and R code for each case study are provided as supporting information (Data S2-4).

TESTING FOR DIVERGENT COMMUNITY DYNAMICS WITHIN A CONTROLLED EXPERIMENT

Porensky et al. (2016a, 2016b) measured plant species composition annually from 2007 to 2014 in quadrats on 24 transects to which six grazing treatments were applied since 1982 (four transects in each treatment): none (exclosure), light, moderate, heavy, heavy-to-light (lightly grazed since 2007; ‘new light’) and heavy-to-none (new exclosure since 2007). They calculated the mean cover in 25 quadrats for each species on transects. We tested the hypothesis that increasing grazing intensity causes increasing divergence (directional change; as detailed in section B) in species composition relative to the control (no-grazing) treatment.

EDA (as described in section D)) showed that the 24 transects were dominated across all measurements by ten common species out of a total of 92 species (Data S2: row and column summaries, relative abundance plots and rank abundance curves). Species turnover (section D 4.0) through time was high; mostly 30% or greater for time period-treatment combinations (Fig. 8). Principal coordinates analysis (PCoA; section D 1.4 Ordination) suggested that the main compositional differences corresponded to the grazing treatments, and that transect composition changed over time, but not markedly (Data S2: PCoA analysis and figure). PERMANOVA (section D 1.4a) showed that transects were significantly compositionally distinct and changed significantly over time in different ways (Data S2: PERMANOVA outputs – all three coefficients for a two-way PERMANOVA with transect,

year and their interaction were significant $P < 0.05$). Partitioning spatial and temporal beta diversity (section D 1.2) showed that two of the control (exclosure) treatments differed significantly (Fig. 9). Principal Response curves (section D 1.4b) showed that the grazed treatments differed from the control and that the relative abundances of four species in particular were influential in driving these differences among treatments, but did not show increasing divergence over time (Fig. 10). This result is consistent with the analysis of temporal coherence ($icc = -0.04$; see section D 1.4a and Data S2 for detailed description of this statistic), which showed that transects were not synchronous in their temporal changes (Data S2: temporal coherence analysis). Thus, we conclude that treatment differences in turnover and shifts in composition were not consistent or directional and so we reject our hypothesis that increasingly grazed transects diverged more over time relative to the control.

TESTING FOR NEUTRAL COMMUNITY DYNAMICS ACROSS SPATIAL SCALES

The 50-ha forest census plot on Barro Colorado Island, Panama, has been censused eight times between 1982 and 2015 (Condit 1998, Hubbell et al. 1999, <http://ctfs.si.edu/webatlas/datasets/bci/>, 2005). This plot was established in part to test for neutral community dynamics (Hubbell 1979) by fully censusing all woody stems greater than 1 cm diameter at breast height every five years. Under this hypothesis, we expected to see no directional or other patterns in temporal community dynamics ('persistence'; as described in section B.). EDA showed that the dataset consisted of 324 species generating a highly right-skewed abundance distribution (most species were rare) (Data S3: EDA analyses and section D). A line graph plot of species' abundance values over time showed very little change in total abundance over time (Data S3: line graph of abundance; Appendix S2: Univariate linear methods). Two analyses that allowed us to test for stability in community composition in time, i.e., change in species composition is lower than expected if individual trees were randomly distributed in time (keeping their spatial positions fixed), showed that change in

species composition was lower than expected: (1) Zeta diversity analysis (Data S3: zeta diversity and section D 2.0) showed that the presence of species across census times was more consistent than expected (Fig. 11) and (2) total turnover in species abundances (Data S3: total turnover rates analyses and section D 4.0) at each of the seven time periods was lower than expected (Fig. 12), particularly when the plot data were at larger spatial scales. Thus, we conclude that the observed lack of pattern is consistent with our hypothesis of neutral community dynamics.

TESTING FOR SEASONAL COMMUNITY DYNAMICS

Gilbert et al. (2012) measured bacterial taxonomic composition monthly over six years using pyrosequencing of the 16S rRNA gene for water samples from a single site off the coast of Plymouth, United Kingdom (source data can be obtained from the journal's online supplementary file). We tested the hypothesis that subsets of the community would show cyclical patterns of temporal change due to seasonality. EDA (section D) showed that the dataset consisted of 74 operational taxonomic units (OTUs) across 62 temporal samples. Most taxa were common through time. Proportional turnover between temporal samples varied from 4% to 48% (Data S4: turnover calculation and section D 4.0), which is relatively low, and was consistent with the detrended correspondence analysis (DCA) result showing a short gradient length of 1.8 for axis 1 (Data S4: DCA and section D. EDA). Similarly, PCoA graphs (section D. EDA) showed little consistent change within years (Fig. 13). However, 13 of the OTUs were deemed to be highly variable in that they changed by more than 500 reads across the sampling period (mean number of reads = 296 ± 829 S.D.; Data S4: row and column summaries, relative abundance distributions and abundance plots and section D. EDA).

Time-lag analysis (section D 1.3) showed that although there was a significant increase in dissimilarity within increasing time lag, variability was high, showing that

temporal changes did not occur in a simple divergence over time pattern (Fig. 14). Temporal synchrony analysis (sensu Loreau and de Mazancourt 2008) showed that some subsets of OTUs in the community were significantly changing in similar ways (Synchrony = 0.27; $P < 0.05$; Data S4: temporal synchrony analysis and section D 3.0). When patterns in these variable taxa were examined through time using zeta diversity analysis (Hui and McGeoch 2014) (Data S4: zeta diversity analysis; section D 2.0), which is based on the presence of species in samples, winter samples shared significantly more species across sets of temporal samples than autumn, spring and summer (Fig. 15; ANOVA, $P < 0.01$; Appendix S2: Univariate linear methods). Further analysis of the most variable taxa using the raw dissimilarity values based on abundance data calculated for temporal samples within seasons showed clear cyclical patterns (Fig. 16; Data S4; section D 1.1). Thus, we conclude that the patterns are consistent with our hypothesis of seasonal dynamics in community composition, largely driven by a small subset of OTUs that varied in their abundance, rather than presence, over time.

D. ROAD MAP FOR ANALYSING TEMPORAL COMMUNITY DYNAMICS

The two most common sets of methods seen in our review of the temporal community dynamics literature were descriptive (mostly tables and figures that show changes in taxon presence or abundance) and ordination-based methods. Although these techniques are useful for illustrating temporal trends and qualitatively exploring patterns in the data, they are used less often for testing specific quantitative hypotheses of patterns and processes in community dynamics. We recommend that researchers augment such descriptions with methods that quantify temporal dynamics in taxonomic composition in more powerful ways, thus allowing more quantitative and potentially comparable descriptions of spatiotemporal patterns and explicit hypothesis-testing of mechanisms underlying observed patterns in community dynamics.

Once the hypotheses to be tested have been generated, we suggest that researchers undertake five steps: (Step 1) perform exploratory data analysis (EDA), (Step 2) choose multivariate methods that can test the *a priori* temporal hypotheses, (Step 3) test for statistical significance, (Step 4) report on the analysis workflow, and (Step 5) consider how to present the results of the selected methods in the text and in figures.

STEP 1: EXPLORATORY DATA ANALYSIS (EDA)

As with any data analysis, the first step in the analysis of temporal community dynamics is an exploratory data analysis (EDA) phase (Zuur et al. 2010). This generates basic information that can be used to both error-check the data and to gain an understanding of key aspects of the dataset (Table 1). R code for the relevant EDA steps is provided in three worked examples, described previously (Data S2-4).

a. Count the number of taxa, the number of samples and the number of empty samples and missing data points

The ratio of the number of taxa and samples is an important consideration because some multivariate methods are more susceptible than others to the influence of rare taxa (Lepš and Šmilauer 2003) and these may be of more or less interest, depending on the research hypotheses. Whilst long-term data and monitoring are highly valued (e.g., Hobbs et al. 2007), it is unfortunately common for long-term datasets to contain missing values; these values should be denoted as missing (i.e., ‘NA’ in R) and not given zero or other numerical values. The importance of missing values and how to deal with them during data analysis depend on the number of missing data points and how they are distributed across the dataset. Multivariate analyses based on compositional dissimilarities, e.g., some ordination methods, cannot be applied to datasets containing samples where no taxa are present (McCune and Grace 2002). However, empty samples may be important for analyses where we are

interested in understanding responses to disturbance and local extinctions. The ratio of the number of explanatory variables to the number of independent replicates is also important for temporal analyses because it affects the degrees of freedom available for statistical testing (Johnstone and Titterton 2009).

b. Examine taxon frequency distributions and response curves

Some multivariate analyses assume specific frequency distributions of taxa (a histogram of the number of individuals or their relative abundance across samples for each taxon), e.g., CCA assumes that taxa have optima along environmental gradients ('unimodal' responses). Examining the shapes of taxon frequency distributions is necessary for determining whether taxa are responding strongly to environmental gradients in the dataset and whether their distributions conform to analysis assumptions. Where many taxa are involved, the common ones should be examined.

c. Determine the ratio of temporal to spatial variation in composition

The degree of spatial and temporal replication is important in determining the method of analysis because few sites with many measurements over time would be analysed differently from many sites measured only a few times. If the data have a nested structure, e.g., samples are nested within sites, this needs to be accounted for in the analysis, particularly if one wants to report significance values. For example, if plant species composition was recorded in quadrats that were nested within transects that were nested within sites, then quadrats are not fully independent replicates. Similarly, information from technical replicates (such as PCR products) should be kept separate at this stage because they provide useful information about variability at different scales in the dataset. It is inadvisable to average or pool data until the significant sources of variation are understood. For instance, if we want to test hypotheses of temporal dynamics, but there is a high degree of spatial variation among samples in the

dataset, this can alter analysis choices or how results are presented and interpreted (Thrush et al. 2008). The relative amount of change in taxonomic composition can be estimated from the differences in sample scores on DCA ordination axes, e.g., the ‘gradient length’ statistic quantifies the turnover in taxonomic composition (Lepš and Šmilauer 2003); this relatively descriptive analysis would ideally be conducted as part of the EDA process prior to selecting an analysis pathway appropriate for testing a quantitative hypothesis of temporal dynamics, such as divergence or convergence.

d. Determine the optimal dissimilarity measure

A community dataset may be an abundance matrix (number of individuals), relative abundance matrix (e.g., percent cover, sequence abundances), rank-abundance matrix (ordinal abundance ranks), or a taxon-occurrence (presence-absence) matrix. These different response variables can influence the choice of analysis methods, including which distance measure to use to calculate pairwise community dissimilarity values for samples (see Magurran 2004, or Legendre and Legendre 2012 for a comprehensive description). To ensure that abundance data from different taxa are comparable using dissimilarity, data should be standardised prior to analysis if they are measured on different scales (Legendre and Legendre 2012). For example, taxon abundance values within each sample can be divided by the total abundance within the sample, or can be standardised across all samples so that each taxon has a mean of zero and standard deviation of one (Borcard et al. 2011). Alternatively, a distance measure that incorporates a standardisation can be used, e.g., Hellinger distance (Legendre and Gallagher 2001, Anderson et al. 2011, Legendre and Legendre 2012). The commonly-used Euclidean distance measure is not usually appropriate for taxon abundance data, which are often characterised by non-linear relationships among taxa, abundance values measured on different scales, or by the ‘double-zero problem’, which is the idea that samples that are missing the same taxa and therefore, have lower dissimilarity when using this

distance measure (Borcard et al. 2011). Comparing results based on presence-absence data with those generated using relative abundance data allows us to evaluate the role of taxon abundances in generating temporal patterns (e.g., Angeler and Johnson 2012). Estimating abundances of each population within the community has its own important considerations, i.e., variation in taxon detection probabilities, which have been addressed elsewhere (e.g., animals: Gaston and McArdle 1994, metagenomics data: Amend et al. 2010). Many distance measures cannot be computed if there are empty samples or missing values. Deleting entire rows or estimating missing data so that the dataset can be analysed may be an option in some cases (McCune and Grace 2002). However, other measures, such as the Gower coefficient, can be calculated while ignoring comparisons among samples with missing data (Legendre and Legendre 2012; coefficient S15: 278).

STEP 2: CHOOSE ANALYSES FOR TESTING HYPOTHESES OF TEMPORAL DYNAMICS

In this section, we describe a range of methods for testing hypotheses of community dynamics, including their outputs, highlighting in each case, the relevant ecological questions and the kinds of hypotheses that those analyses can be used to address: pairwise dissimilarity-based methods (section 1, below), zeta diversity (section 2), synchrony (section 3), turnover rates (section 4) and multivariate regression modelling (section 5). Further, each of the described methods has been cross-referenced with (i) which hypotheses it is best suited to test; (ii) whether or not it allows the identification of individual taxon effects on, or responses to, temporal change; (iii) whether or not explanatory and spatial variables can be explicitly incorporated; (iv) the type of statistical testing normally used; and (v) any requirements for the length of the time series (number of temporal samples) (Fig. 17). The selection of an appropriate analytical method also can be influenced by the kind of output desired (Table 2). For instance, some analyses result in outputs that require more effort to interpret than others, whilst others show differences in taxon responses or generate outputs that can be added to a

geographic map. Further, as our literature review highlighted, in many cases several different analytical methods often are appropriate for asking questions about temporal dynamics and presenting the results at the desired level of complexity; it is also the case that many analyses can be used to test a wide range of hypotheses, so our focus is on describing the strengths of each analysis method and the minimum requirements for use, rather than on making explicit recommendations. Thus, we recommend that, at this point, readers use Tables 2 and 3 to help identify the methods of interest and then refer to the relevant sub-sections of section C that describe those particular methods.

Our literature review revealed that in numerous cases, researchers performed and compared separate analyses on their community matrices from each sample time, e.g., separate nestedness analyses at each sample time (Azeria et al. 2006); we do not specifically address these applications here in any depth because these methods do not explicitly evaluate community dynamics through time. We also focus on methods that are truly multivariate rather than methods that are applied separately to abundances of individual taxa; however, if less than ten taxa are of interest, it may be that hypothesis testing via univariate approaches (e.g., separate linear regressions of taxon abundances) will be adequate for drawing synthetic conclusions about community change. In the supplementary material (Appendix S2), we summarise univariate methods that were commonly used in the literature review articles: univariate linear methods, temporal stability (*a.k.a.* coefficient of variation), frequency change, multiplicative change, and Markov chain modelling. A few studies used alternative methods such as cluster analysis and network analysis; however, we excluded those because we think the methods presented here provide equally good, or better, alternatives, but we nonetheless provide an explanation of them (Appendix S2) because they appeared in our literature review. Some of the analytical methods described below can be implemented in several different software packages (e.g., PC-ORD, PRIMER, CANOCO), but, for simplicity

and widest accessibility, we illustrate their implementation in R (Data S2-4), except for Moran's Eigenvector Maps (MEMs), which is given in Legendre & Gauthier (2014), and Multivariate Autoregressive State Space modelling (MARSS), which is presented in the documentation associated with the R package 'MARSS' (Holmes et al. 2012).

1. Methods based on pairwise dissimilarity values

Pairwise dissimilarity methods can be used to test nearly the full range of questions and hypotheses in community dynamics (Anderson et al. 2011; Figs. 4-7). Pairwise dissimilarity methods are the most frequent way ecologists quantify the difference in the identities of taxa occurring in two samples. The result is a single number that measures the distance in compositional space between the two samples. Dissimilarity can also be represented as similarity, i.e., $1 - \text{dissimilarity}$ (Whittaker 1960, Anderson et al. 2011, Magurran 2013). When applied to multiple samples, it results in a matrix of all possible comparisons (a triangular dissimilarity matrix). For example, Collins (1992) measured the between-year percent change in plant species composition for grassland plots that had been repeatedly measured over a nine-year period. Dissimilarity methods are more desirable than univariate methods, which analyse the abundance of each taxon separately, because they consider the entire community at once and directly quantify change in community structure.

The development of dissimilarity indices is an active area of research. Magurran (2013) is a useful general discussion of dissimilarity measures, Anderson et al. (2011) provides a practical guide to selecting an appropriate index and issues of importance, such as whether to include joint absences in the calculation. Different authors often use different names for the same dissimilarity measure, so we recommend using a comprehensive and consistent guide, such as Legendre and Legendre (2012; Table 7.4: 324). Legendre and De Cáceres (2013) also provide comprehensive descriptions of most dissimilarity measures and

rationales for choosing one. Several recent advances are specifically relevant to temporal data (e.g., Podani and Schmera 2011, Baselga 2012, Shimadzu et al. 2015) and some contributions aid in improving dissimilarity measures for use in certain circumstances, such as imperfect detection of taxa (e.g., Chao et al. 2006) or which indices are most appropriate for calculating dissimilarity (Legendre and De Cáceres 2013). For example, the unique fraction (UniFrac) dissimilarity metric was designed for microbial communities where phylogenetic information is available from sequencing that has been used to determine taxa in samples (Lozupone and Knight 2005, Lozupone et al. 2007). This metric measures phylogenetic distance between samples as the fraction of branch lengths leading to descendants from each sample. UniFrac can be used with abundance data, which weights the branch lengths prior to calculating distances (weighted UniFrac), or with presence-absence data (unweighted UniFrac Lozupone et al. 2007).

In the context of temporal community dynamics, dissimilarity is calculated in two different ways: (1) separately among spatial samples at each of a series of time points or (2) among samples taken at different times at the same location(s). Dissimilarity matrices are the basis for several methods recommended in our guide, including ordination analyses and Moran's Eigenvector Maps (MEMs). It is important to realise that it is not easy to use such methods to simultaneously assess both temporal and spatial variation in the same analysis; however, temporal and spatial dissimilarity can be computed and assessed in parallel to test hypotheses and make inferences about spatiotemporal community dynamics (e.g., Legendre and Gauthier 2014, Lamy et al. 2015).

1.1. Using the raw dissimilarity values

There are two ways in which raw dissimilarity values for samples are typically used. Either of these approaches could be used for different taxonomic subsets of the community to test hypotheses of taxonomic variation in dynamics. First, compositional dissimilarity can be

calculated among samples taken at different locations to obtain a measure of spatial heterogeneity in community composition. To use these values to analyse temporal community dynamics, pairwise dissimilarity (spatial dissimilarity) must be calculated separately for each sampling time, and the means of these spatial dissimilarity values can then be compared across time visually in a table (e.g., Collins 1992), graph (e.g., Barros et al. 2014), or using a correlation or regression analysis where time is the predictor variable and amount of spatial variability at each site is the response (e.g., Barros et al. 2014). The mean dissimilarity can be divided by the number of taxa to generate an independent measure of compositional change (Cleland et al. 2013). Such comparisons over time could be done separately for different groups of samples, such as different experimental treatments, to test for predictors of differences in temporal change in composition. Bagchi et al. (2017) illustrate predictions of how temporal differences in dissimilarity can include non-linear patterns and thus how applying linear and non-linear regression modelling to pairwise dissimilarity values can be used to test more complex hypotheses of temporal dynamics. Regression approaches can also include mixed-effects models that account for nesting and other types of non-independence in the data, or non-linear models (e.g., generalised additive models) that may be useful to predict cyclical or threshold patterns in community dynamics (e.g., Shenhav et al. 2017). However, these modelling methods will apply well only in situations with longer time series (i.e., greater than 25 temporal replicates) and the more complex the prediction of temporal change, the greater replication is likely required. Alternatively, a Mantel test can be used to compare pairs of dissimilarity matrices computed at different times and to test for significant difference over time (see section in the main text ‘Test for statistical significance’). This approach is only useful for a small number of spatiotemporal observations (i.e., ten or fewer) because the comparisons are pairwise.

Second, if a set of samples were all taken at the same site at different times, ‘temporal heterogeneity’ can be calculated using the mean dissimilarity of the entire sequence of temporal samples to obtain a single measure of within-site temporal variability (Collins and Smith 2006). If multiple sites are sampled through time, this method can be applied at different spatial scales (Collins 1992, Collins and Xia 2015) to examine spatial scale-dependent temporal dynamics (Collins and Smith 2006). Regression approaches that relate mean temporal dissimilarity values to other explanatory variables for sites or samples could be used to explore spatial variability in temporal change. Alternatively, the ‘dissimilarity-overlap’ method, first described by Bashan et al. (2016) and further explained by Kalyuzhny and Shnerb (2017), can be applied to raw dissimilarity values to test for similarity among samples in community dynamics. This method uses the relationship between the pairwise dissimilarity values and the proportion of individuals in pairs of samples that overlap taxonomically to detect if changes in taxonomic composition (degree of overlap) alters relative abundances in a consistent way across samples (dissimilarities). It can be used for very large numbers of samples and taxa and could be applied to different groups of taxa to test for synchronicity.

1.2. *Beta diversity decomposition*

‘Beta diversity’ is term used for several different types of measures in ecology, e.g., (Whittaker 1960, Baselga 2010, Legendre and De Cáceres 2013), including as a synonym for compositional dissimilarity. It always describes some measure of how taxonomic composition changes across a set of samples. A number of recent advances in the calculation of beta diversity have illustrated that, if calculated in particular ways, it can be ‘partitioned’ into different components that relate to different predictions of how taxa or their abundances change in space or time (Baselga 2010, Ellison 2010, Legendre and De Cáceres 2013). Beta diversity decomposition methods can be applied to testing almost any hypothesis of

compositional change (Fig. 17). The outputs are focussed on species replacement and whether samples are subsets of each other (Table 2), so detecting patterns such as convergence or divergence is very easy. However, depending on what subsets of the data are used in separate calculations, beta diversity decomposition is equally able to detect rapid shifts in composition, e.g., perturbation, tipping point, or increases or decreases in temporal variability, e.g., ‘instability’ or ‘stability’. Two approaches so far have been usefully applied to quantifying community dynamics measured in the field (Baselga 2010, Baselga and Orme 2012, Legendre and De Cáceres 2013, Lamy et al. 2015).

First, when among-sample dissimilarity is measured as Sørensen’s dissimilarity $[(b + c)/2a + b + c]$, where a is the number of shared presences and b and c are the number of unshared presences in each of the two samples], the Simpson’s dissimilarity index (also known as ‘species turnover’: $b / (b + a)$) can be subtracted from this value, resulting in a measure of ‘nestedness’ (Baselga 2010, 2012, Baselga and Orme 2012). This measure of nestedness reflects the degree to which the species composition of smaller samples (lower species richness) is encompassed by the species composition of larger samples (higher richness); samples may differ in space or time. In contrast, ‘species turnover’ measures the degree to which samples do not share species: how the identities of species differ among samples. It is useful to separate these two components (nestedness and turnover) to obtain a more informative quantification of beta diversity that reflects differences in the underlying processes. Such separation can be done either pairwise or for multiple sets of samples (Baselga 2010, 2017), and has been used to decompose temporal beta diversity in tests of hypotheses unique to temporal data (e.g., Baselga and Orme 2012).

Second, and in contrast, Legendre & De Cáceres (2013) defined beta diversity as the sums of squares in the variation of taxon abundance (or presence) corrected for the number of samples. They partitioned this measure of beta diversity into two different components:

‘Species contribution to beta diversity’ (SCBD) and ‘local contribution to beta diversity’ (LCBD), which describe, respectively, the relative contributions of different taxa and sites to overall beta diversity (Table 2; Legendre and Gauthier 2014). This method allows for tests of predictions of patterns in temporal community dynamics related to different taxa in the dataset. For example, Lamy et al. (2015) decomposed beta diversity (based on Bray-Curtis dissimilarity) of fish communities at each of 13 reef sites into contributions from biomass replacement and biomass differences over time. They demonstrated that total biomass in the community changed little, but that the observed significant temporal changes in fish composition resulted from differential changes in the relative biomass of species. Spatiotemporal patterns in the two components of fish community beta diversity were visualised using principal coordinates analysis (PCoA).

1.3. Time-lag analysis

A time-lag analysis involves relating the amount change in composition to the amount of change in time across increasing temporal distances, called ‘lags’ (Table 2). For example, Collins & Xia (Collins and Xia 2015) compared plant species composition of repeatedly-measured grassland transects across all time lags between 1989-2008 (from a 1- to 19-year difference). There are two ways of assessing the time-lag effect on community dissimilarity: graphically (‘time-decay curve’) and statistically (time-lag regression analysis). For meaningful inference, time-lag graphs require at least five samples and statistical time-lag analysis requires at least ten samples for each temporal sequence. For both approaches, the pairwise compositional dissimilarity (temporal beta diversity) among temporal samples is calculated first.

To generate a time-decay curve, pairwise temporal dissimilarity values are plotted against temporal distance between samples (time lag; pairwise Euclidean distances in sample times) to show how community dissimilarity changes as samples become further apart in time

(Collins 2000, Collins et al. 2000, Hallett et al. 2016). This graph can be then used to visually assess whether community composition is converging (becoming more similar) or diverging (becoming less similar) over time (e.g., Collins 2000). For example, Bar-Massada & Hadar (2017) plotted temporal beta diversity (as Bray-Curtis dissimilarity) of plants in grazed and ungrazed treatments in Mediterranean herb communities, revealing that composition diverged over time, regardless of grazing treatment. Alternatively, temporal community dissimilarity can be graphed against dissimilarity in environmental variables (rather than time), if the primary research focus is on response of the community to environmental change rather than time *per se*.

Time-lag regression analysis was developed specifically for data with greater than ten time points to assess directional, cyclic or stochastic changes within communities (Collins 2000, Collins et al. 2000) and has been evaluated against other temporal analysis methods (Angeler et al. 2009). For time-lag regression analysis, the pairwise compositional distances are plotted against the square-root of the time lag, which reduces the bias that results from having fewer data points at larger time lags (Collins 2000, Collins et al. 2000). A regression line is then fitted to determine the pattern of compositional change. A significantly positive relationship indicates that compositional changes in communities is directional, because sampling points that are further apart in time are more dissimilar than those that are closer together in time (i.e., Fig. 2 in Collins 2000). A negative slope shows that some convergence towards the taxonomic composition at one of the earlier time points has occurred, because measurements close together in time are more dissimilar in composition compared to those further apart in time. No slope, or a non-significant regression, implies stochastic changes in the community over time. Non-linear regressions also can be used to describe more complex temporal patterns (Magalhães et al. 2007).

A robust measure of significance in time-lag regression analysis can be determined by running Mantel tests on the dissimilarity matrix and the matrix of pairwise distances among samples in time, with Monte Carlo permutations (Bêche and Resh 2007, Day et al. 2017). Time-lag regression analysis is appealing because it is intuitive, produces one statistic (the regression coefficient), and can be easily visualised (e.g., freshwater macroinvertebrates: Bêche and Resh 2007, gut bacteria: Dimitriu et al. 2013, and plants: Collins and Xia 2015). However, time-lag analysis only provides one value for each sampling unit, so further analyses may be required to investigate changes within treatments or groups.

1.4. Ordination

Ordination summarises multivariate community data by optimising relationships between high-dimensional samples and taxa in low-dimensional space to detect important ecological gradients in communities (Table 2). This dimensionality reduction means that they are not especially useful for directly testing, or visualising, complex changes in community composition over time, such as tipping points, cyclical patterns or synchronicity. However, more simple patterns, such as divergence or convergence, can be more easily detected. Ordinations are either unconstrained, meaning that they do not incorporate explanatory variables, or constrained, which assess the effect of explanatory variables, including time (e.g., Kent 2012, Legendre and Legendre 2012, Palmer 2019). Ordination methods, including constrained ordination and using the ordination scores from an unconstrained ordination in a regression model (e.g., general linear model, mixed-effects model) are only appropriate for testing hypotheses of community dynamics in cases where EDA shows that the ordination axes used strongly represent temporal compositional change and therefore explain a substantial proportion of the variation. This is because the amount of temporal compositional variation represented by the ordination axis scores will be low if there is not a strong temporal gradient along the axes. Ordination methods allow only weak inference of the

relationship between taxa and community dynamics if the temporal pattern is complex relative to the ordination axes because some ordinations, e.g., DCA, only provide an ordination axis score for taxa, and others do not provide ‘species scores’ at all, e.g., PCoA and NMDS. PCoA can be used to calculate the amount of variance explained by the ordination axes, which is important if a time-series analysis is to be performed using the axis scores. Where species scores are not provided, weighted averaging can be used to plot points for each taxon on the ordination diagrams to relate differences in site composition to differences in taxon presence or relative abundance if you wish to examine the associations between samples and taxa (McCune and Grace 2002). One advantage of PCoA is that the first PCoA axis represents the gradient explaining the greatest amount of variation in composition, followed by the second axis that explains the next greatest amount, and so on; this is not the case for NMDS (McCune and Grace 2002). Therefore, it is recommended that NMDS scores are not used in community dynamics analyses unless the amount of variation explained by each axis is clearly stated in order to validate their use in further analyses. ‘Co-inertia analysis’ can be coupled with ordination in a ‘partial triadic analysis’ to investigate community dynamics in relation to environmental variables. This method may be particularly useful when testing for temporal stability (Kidé et al. 2015).

a. Using ordination site scores to assess compositional change

Ordination axis scores for samples taken at different times can be used to quantify and test hypotheses of compositional change. Ordination diagrams are visually appealing, but for illustrating compositional change through time, it is difficult to interpret the magnitude and significance of temporal dynamics without statistical testing, such as permutational analysis of variance (PERMANOVA: Anderson 2001), unless temporal trends in the data are very strong along the ordination gradients in composition. This difficulty is amplified if there are many samples or if the sampling design includes a large number of sites, environmental

factors, or temporal replicates. Note that PERMANOVA requires the assumption of homogeneity of dispersion to be met (this can be tested using the “betadisper” function in the R package *vegan*; Oksanen et al. 2017). Further, ordination diagrams and their statistics computed for different matrices (different sets of taxa or sites) are not comparable; results of individual ordinations cannot be compared among them. The partial exception to this is Procrustes analysis, which can measure the distortion, or ‘lack of fit’ between two ordinations, or dissimilarity matrices (Jackson 1995, Peres-Neto and Jackson 2001). It can be used to identify which taxa are most important in differentiating between the two ordination patterns.

A more quantitative, and often more flexible, approach is to use the ordination axis ‘site’ scores in linear models with time and other explanatory variables of interest entering as predictor variables (e.g., Berthon et al. 2014, Broadway et al. 2015). For example, Cotton et al. (2015) used DCA scores as a response variable in a linear mixed model to investigate the importance of carbon dioxide, ozone, and time on arbuscular mycorrhizal fungal community composition. This is like doing a constrained ordination, such as RDA, dbRDA, or CCA, where time is used as a constraining explanatory variable in the ordination; these methods will test the significance of time and any other explanatory variables. However, it is more difficult within a constrained ordination to test for non-linear, cyclical or other complex temporal patterns, and in this case, using the ordination axis scores in a more complex model may be more useful.

Ordination axis scores for samples can be used to quantify temporal changes in other ways. For example, Quinn et al. (1991) developed a ‘malleability index’ by regressing the pairwise Euclidean distances between the first three dimensions of ordination space for temporal samples, against environmental variables. One potential problem with this method is caused by the tendency of ordinations to distort the relative distances (dissimilarity)

between samples, depending on, for example, differences in the number of rare taxa, and so values from distance-based indices may not be comparable, even in different parts of the same ordination.

Temporal coherence measures the degree to which a response variable is similarly variable in time across a set of sites. In the case of measuring temporal coherence in community composition, the multivariate response variable is taxonomic composition in the form of ordination axis site scores (Angeler and Johnson 2012). Temporal coherence compares the amount of temporal variation in the ordination axis site scores across a set of sites by using statistics generated by an ANOVA of the site scores (using year and site as factors) to calculate the ‘intra-class correlation coefficient’, which ranges from $-1/(n-1)$ to $+1$, where n is the number of sites sampled, given that $n > 1$ (for calculation details see Zar 1996 pp 398, Rusak et al. 1999). High positive values of this metric indicate increased coherence of the different sites in composition over time. The measure can be recalculated while dropping each time in sequence (an ordered jackknife) to determine the relative effect of each time on the overall value (Rusak et al. 1999). Different sets of sites or treatments can be compared using a Fisher’s Z test (Angeler and Johnson 2012) or randomisation test in the case of low sample sizes (Rusak et al. 1999).

b. Principal Response Curves (PRC)

The method of principal response curves (PRC) was first developed by van den Brink & ter Braak (1999) for comparing compositional change in treatment sites relative to a control site. It is similar to redundancy analysis (RDA) in that a multivariate regression is performed on the sample $\times$ taxon matrix constrained by variables of interest, but PRC is explicitly constrained by time (Lepš and Šmilauer 2003, ter Braak and Šmilauer 2015). It was developed to eliminate the need for cluttered biplots that make it difficult to see temporal trends within and between treatments. To overcome this problem, PRC uses a single axis,

plotted against time, to express trajectories of compositional change (Lepš and Šmilauer 2003: 145); further axes can be calculated if they are of interest. Taxon weights also can be calculated and plotted for each PRC to show the change in the relative abundance (e.g., percent cover) of each taxon in the treatments relative to the control. The significance of the PRC is tested using Monte Carlo simulations (Van den Brink and Ter Braak 1999). As with RDA, PRC yields a measure of the percent variance explained for the overall taxonomic composition and how much is attributable to each variable (treatment and time). For example, Sugiura et al. (2013) used PRC to show that plant invasion in forests on a small island in the north-western Pacific had little effect on the temporal dynamics of invertebrate communities. The potential drawbacks of this method include its requirement for a ‘baseline’ or ‘control’ treatment level to be set for comparison, an attribute that many observational datasets will not have. Furthermore, PRC allows for only one ‘treatment’ variable and each analysis focusses on the change at one site, relative to the control site. However, Auber et al. (2017) have recently generalised this method so that it can be used to analyse compositional change at multiple sites, relative to a baseline or control, but this application removes the ability to consider an environmental or treatment effect. Thus, a principal response curve analysis yields a line graph showing the relative change in taxonomic composition for each ‘treatment’ along one compositional gradient relative to a control (Table 2).

1.5. *Moran’s eigenvector maps (MEMs)*

Moran’s eigenvector maps (MEMs; also known as ‘principal coordinates of neighbour matrices’, PCNM) is a method designed to characterise the spatial scales of variation in univariate or multivariate data (Borcard and Legendre 2002, Dray et al. 2006, Legendre and Gauthier 2014). However, MEMs have recently been applied to detect temporal scales of variation in communities (Angeler et al. 2009, Caruso et al. 2013, Legendre and Gauthier 2014). For instance, if a cyclical pattern in compositional change, such as seasonality, was

detected as significant for a dataset, this can be illustrated using line graphs showing the temporal pattern (Table 3; Legendre and Gauthier 2014). This method is particularly useful for detecting very complex temporal patterns, such as seasonality (periodicity), cyclical changes or tipping points, but requires relatively large temporal sample sizes (Fig. 17). For example, Kampichler et al. (2014) used MEMs to detect temporal patterns in multiple datasets that had at least 20 annual measurements of bird communities.

First, a distance matrix is generated. Then, to focus on neighbour relationships, a threshold value is chosen (see Borcard et al. 2011 for advice). For values below this threshold, the distance value is retained; values above this threshold are assigned an arbitrarily large value. Legendre & Gauthier (2014) recommend using one time interval as the threshold for regular time series or the length of the longest lag for irregular series. A PCoA is then applied to this modified distance matrix to generate a large number of eigenvectors that describe a wide range of temporal scales of variation.

The eigenvectors from the PCoA can be used as explanatory variables, most often in a redundancy analysis (RDA) or multiple regression. Combined with RDA, MEMs can be used to determine whether there is cyclic or synchronised temporal change in composition and which taxa, or taxon groups, are the drivers of it (Fig. 17). MEMs-RDA scores also can be plotted as a function of time to assess visually the intervals of fluctuating community dynamics (Angeler et al. 2009, Kampichler et al. 2014). In a comparison of different methods to assess community dynamics, Angeler et al. (2009) observed that MEMs-RDA consistently explained a greater proportion of variation in data with low taxonomic richness and linear changes in time relative to time-lag regression analysis, Mantel tests, or RDA. However, power was higher for taxon-rich communities because there was a greater probability of more taxa fitting the temporal models. As a scale-detection method, MEMs could be applied to determining the relevant time intervals for community monitoring. This method can be

combined with variance partitioning to show the relative variance in composition explained by temporal, spatial and other explanatory variables (Legendre and Gauthier 2014).

1.6. Compositional pivot days

Compositional pivot days is a dissimilarity-based method developed by, and unique to, Lellouch et al. (2014). This method is appropriate for situations where we want to test for the occurrence of fast changes in taxonomic composition across one point in time, such as the application of an experimental treatment or the effects of a rapid change in environmental conditions (Table 2). It works by comparing pairwise distance values for pairs of temporal observations taken at the same sample location to identify the time points where rapid shifts in composition occurred. A ‘pivot day’ is a sampled time point that is identified by pairwise similarities being low before and after that time point, but relatively higher across it. Compositional pivot days are identified either visually from the pairwise distance matrix plot, or using clustering (see Lellouch et al. 2014 for details). This method may be a good option for a preliminary identification of thresholds or rapid shifts in community structure, but does not provide a statistical test for the significance of the threshold relative to background. The power of this analysis will increase as the number of sampled time points increases relative to the number of pivot days.

1.7. Community trajectory analysis

Community trajectory analysis quantitatively compares the temporal change in multivariate compositional dissimilarity space among samples using geometric calculations (de Caceres et al. (2019)). These comparisons are based on the relative amounts, rates, and directions of sample movement in compositional space, thus, the method allows flexible testing of a wide range of very specific, quantitative hypotheses of change, such as the divergence or convergence of samples over time. This method is especially useful for a small number of

spatial replicates because the output can be graphical; plots use arrows on an ordination diagram to show how samples have moved in compositional space over time; however, larger numbers of samples could be graphed separately. We do not illustrate this method here because comprehensive R code and explanations are presented in the R package ‘vegeclust’ (De Cáceres et al. 2010) and demonstrative examples are given in de Cáceres et al. (2019).

2. Zeta diversity

Zeta diversity, a measure of similarity, analyses change in community composition among samples in the dataset by calculating the average number of taxa that are shared between samples that are grouped in pairs, triples, quadruples, etc. (Table 3). The method was developed and introduced by Hui & McGeoch (2014), extended to temporal data by McGeoch et al. (2017), and comes with its own R package, ‘zetadiv’ (Latombe et al. 2017b). Zeta diversity estimates compositional turnover from multiple ($n \geq 2$) samples, rather than the pairwise comparisons used to estimate beta diversity. What is obtained by the method is a series of zeta values (similarity) for different orders of zeta: zeta_1 to zeta_n , where n is the total number of time points. Each of these values represents the average of the number of taxa shared by each set of times in the dataset. For example, zeta_2 (zeta order 2) represents the average number of taxa shared by all possible pairs of samples in the dataset (equivalent to 1 – beta diversity or mean similarity), zeta_3 represents the average number of taxa shared by all possible subsets of three samples, and so on. Zeta_1 (zeta order 1) represents the average alpha diversity (taxon richness) of all the samples in the dataset.

If there is complete turnover in community composition in time, at some point, zeta_i (where $i \neq 1$) will equal zero when there are no taxa shared by a given sized subset of samples. In contrast, if community composition never changes through time, $\text{zeta}_n = \text{zeta}_1$. Therefore, the rate of decline in zeta with increasing ‘zeta order’ tells us about the rate of

change in community composition in time (‘zeta decline’). ‘Zeta decay’ can be considered analogous to time-lag graphs and time-lag regression analyses, which we described earlier for pairwise dissimilarity measures (Table 3). This means it experiences the same issue of few time points at higher lags, and thus higher uncertainty. The temporal scale(s) of community dynamics can be illustrated by plotting the change in sample zeta diversity of a particular zeta order against the temporal lag of samples, meaning that this method is useful for detecting both directional changes and cyclical patterns. For instance, if a multi-year time series in composition is cyclical (e.g., shows seasonality), zeta would seasonally decrease and then increase. These analyses can be done for different sets of samples to make spatial or temporal comparisons and, if necessary, zeta values can be normalised using the gamma diversity (total number of taxa in the dataset) to account for differences in taxon richness across samples (McGeoch et al. 2017). Further, different methods of creating the subsets for calculating zeta can be used, such as nearest neighbours or relative to a time-zero sample, further facilitating hypothesis testing (McGeoch et al. 2017). Latombe et al. (2017a) used zeta diversity to extend generalised dissimilarity modelling (Ferrier et al. 2007). This enables the user to assess taxonomic turnover at different orders of zeta and understand drivers of turnover in rare and common taxa (Latombe et al. 2017a). This method could be extended to assess temporal turnover at different orders of zeta by adding time as an environmental variable.

3. Synchrony

Although other methods can be performed on taxonomic subsets to look for similarities and differences in the temporal dynamics among groups of taxa (“synchronicity”), a useful measure of the ‘temporal synchrony’ in multivariate taxon abundances was developed by Loreau & Mazancourt (2008) and comes with its own R package, ‘synchrony’ (Gouhier and Guichard 2014). This standardised value measures how closely changes in taxon abundances

within a community track one another in time. It varies between zero (complete asynchrony of taxon abundances in the community) and one (perfect synchrony; all taxon abundances changing in the same way) and, by exploring different sized ‘time windows’, can be used to detect scales of temporal variation. Overall significance tests of synchronicity in community change can be obtained by performing null model analysis (e.g., Pedersen et al. 2017).

A similar method (Houlahan et al. 2007) uses the sum of the temporal variances and pairwise covariances of all taxa in the community to distinguish if the community shows ‘compensatory dynamics’ over time. This method results in either a positive, negative or zero community-level covariance value, indicating that taxa are increasing or decreasing in synchrony (positive community covariance), some taxa are increasing, but this is compensated for by decreases in other taxa (negative community covariance), or that no change is evidence (zero community covariance) (Houlahan et al. 2007, Hallett et al. 2016).

4. Turnover Rates

The degree of temporal variation in community composition can be assessed by calculating the turnover rate of the community (a.k.a. ‘temporal turnover’ or ‘species turnover’), which is a measure of the rate of change in taxonomic composition for a ‘site’ over time (e.g., Aguirre et al. 2003). Turnover rates can be computed to measure the magnitude of compositional change, but in a way that does not result in an understanding of the direction of change or the taxa involved in compositional shifts. This method can detect when, and how rapidly, temporal shifts occur in a temporal sequence to detect complex temporal patterns such as perturbations, tipping points and pivot points (Fig. 17). Turnover rates are calculated using combinations of colonisation, immigration, extinction, mortality, recruitment, and survival. One of the commonly used methods is that of Diamond (1969) where the percentage turnover rate, T , is calculated as $T = 100 \times (E + C) / (S1 + S2)$, where E is the number of taxa that went

extinct between two time points, C is the number of taxa that colonised between the same two time points, S_1 is the number of taxa present at time 1 and S_2 is the number of taxa present at time 2. T gives the percent change in taxon identities over the time period and varies from 0, representing no turnover, to 1, indicating complete turnover (Diamond 1969). This measure can be calculated only for a single location at two time points; pairwise turnovers must be calculated separately for datasets with multiple time periods or multiple sites. T values can be compared through time, at different time intervals, or at different spatial scales (e.g., Boulinier et al. 2001, Tworek 2004). They also can be used in a regression-style analysis to assess the effects of a treatment variable or other explanatory factor on compositional turnover (e.g., Aguirre et al. 2003, Tworek 2004). These methods are only useful for communities that have relatively high rates of temporal turnover in taxa. The degree of compositional turnover also is tied to the temporal extent of the study for the given system and the average lifespan of taxa. For example, complete turnover in prokaryotes can be observed over a period of days (e.g., Kara and Shade 2009), whilst many decades may be required to detect any species turnover in some plant communities (e.g., Kardol et al. 2010). However, because it can be standardised by the number of taxa in the community, T is a potentially comparable measure of compositional change among different datasets, if they are on similar timescales and encompass communities with a similar distribution of life-histories.

5. Multivariate regression modelling

Multivariate autoregressive state-space models (MARSS) are multivariate linear regression methods that model the stochastic temporal dynamics in matrices of species' abundances in relation to abiotic or biotic covariates, while accounting for potential observation (measurement) error (Hampton et al. 2013). The models are autoregressive in that previous measurements in time are assumed to be correlated to future measurement, and these

autocorrelations are formalised by modelling the current values of the response matrix as a weighted linear sum of previous values at a specified temporal lag (Penny and Harrison 2006). MARSS models can be used to disentangle temporal variation in species abundances into the relative contributions of inter- and intraspecific interactions, abiotic variations, and of demographic stochasticity (Mutshinda et al. 2009). Effect sizes of these different components are estimated from the multivariate model structure through the use of conditional least-squares, maximum-likelihood, or Bayesian estimation methods (Hampton et al. 2013). Observation errors, and other complexities such as spatially-replicated time series and non-uniform sampling, can be formally modelled within the state-space framework (Holmes et al. 2012). Further, MARSS models can be used to compute several metrics that indicate the degree to which communities are possibly becoming less stable over time (Ives et al. 2003). This level of flexibility means that they are able to test for even the most complex temporal patterns in communities, and can simultaneously reveal similarities and differences among taxa, and account for environmental factors or other covariates. However, they do require relatively long time series (ideally more than 25 temporal replicates), and, like any statistical model, more independent replication is required for more complex models. There are several examples of the use of MARSS models in freshwater and marine ecology, particularly in terms of estimating aquatic food web dynamics (Beisner et al. 2003, Hampton et al. 2008, Thibaut et al. 2012); there have been relatively few applications for terrestrial communities (e.g., Yamamura et al. 2006, Mutshinda et al. 2009). The ‘MARSS’ R package (Holmes et al. 2012) enables the fitting of MARSS models and includes a range of ecological examples, so we have not included example code in this paper.

Multispecies N mixture models and joint species distribution models are other regression methods that can be used to model explicit, quantitative hypotheses of community change, with a focus on interactions among taxa (e.g., Dorazio et al. 2015, Ovaskainen et al.

2017). These methods can allow for useful additions, such as accounting for uncertainty in the detection of taxa; however, due to their complexity and high computational requirements, they have yet to be implemented for datasets with large numbers of species.

STEP 3: TEST FOR STATISTICAL SIGNIFICANCE

Whichever multivariate analysis has been selected, most ecologists interested in community dynamics seek a way to determine if the changes that they see over time in the results of those analyses are greater than, or less than, they would expect by chance alone. Some of the methods presented above have their own built-in significance tests, often based on randomisations of the community matrix. Commonly-used randomisation tests include the Mantel test (Legendre and Legendre 2012) and permutational multivariate analysis of variance (PERMANOVA; Anderson 2001, Legendre and Legendre 2012). Null models, whereby community matrices are randomised under sensible criteria (Gotelli and Graves 1996, Gotelli 2000, 2001, Gotelli and Entsminger 2001, Ulrich and Gotelli 2007, 2010), can be used to test the significance of any observed change in compositional dissimilarity, and can easily be incorporated into analyses described above including zeta diversity. The primary consideration in applying any randomisation test is what assumptions are being made about taxon niches and whether they are valid. If the interest is in testing the effects of interactions among taxa on their co-occurrence, but the randomisation procedure (null model) allows taxa to occur in samples where they are unable to occur for other reasons, such as habitat differences, a difference among samples may be falsely detected (an inflated Type I error rate). Thus, if strong inferences are to be made about compositional differences, we must be sure that the differences among the groups of sites that we are testing are the only difference. Similarly, the sample size needs to be high enough to allow the detection of differences among sets of sites for the given level of compositional variability (Type II error rate).

Schaefer, et al. (2005) introduced a novel method for detecting significant shifts in community composition using randomisations under null models of taxon mean abundances through time. They demonstrate that this approach, which uses both compositional distance (pairwise dissimilarities) and the coefficient of variation (CV), could detect significant dynamics in community composition, even where only small amounts of change had occurred. In this method, temporal changes in communities are assessed by using a compositional distance measure that is selected by the user. This quantifies the multivariate distance between the randomly generated communities (null model randomisations) and observed ‘reference’ samples (e.g., ‘Time 0’ samples, compared to randomisations of ‘Time 1’ samples). If the observed distance is greater than under randomisation, then we can infer statistically significant temporal change. Because this can be done in an ordination context, increases and decreases of particular taxa driving the change can be determined from the ordination plots. Limitations of this method are that it requires a relatively large degree of temporal change in the data compared to spatial variation (i.e., more temporal, than spatial, variation) and it is relatively computationally intensive; however, we provide the R code for a worked example in the supplementary material (Data S2).

STEP 4: REPORT ON THE ANALYSIS WORKFLOW

Our literature review suggested a lack of consistent reporting of critical aspects of the data analysis workflow. Such reporting ensures that the results are robust, represent the variation deemed important for the research aim and are reproducible. The number of taxa, samples, and independent replicates in space and time all should be clearly reported in the Methods or Results section of the paper (see Hurlbert 1984 for guidance on statistical independence). The dissimilarity measure selected, and the rationale and procedures used to arrive at that decision, should be reported in the Methods section of the paper. We encourage researchers to refer to a specific source, such as Legendre & Legendre (2012), or to give the

formula applied to maximise the repeatability and transparency of their research; these indices can be inherent in particular methods or built-in as defaults in software, but nevertheless should still be reported. Different methods of dissimilarity or ordination (or other analysis types) that were compared, e.g., in the EDA phase, should be listed in the Methods section of the paper stating that their results were consistent with each other. The choice of dissimilarity or ordination method is subjective, but rules of thumb can be used for assessing different methods, so that the method selected optimises the reduction of dimensionality in the data and thus maximises the variation accounted for by the analysis (see McCune and Grace 2002 for comprehensive advice).

STEP 5: PRESENT TEMPORAL PATTERNS AND RESULTS

Some of the analysis methods that we describe in this paper do not produce simple visual outputs that clearly show temporal trends in community datasets and visualising temporal dynamics can be difficult for species-rich datasets or complex sampling designs. However, there are a variety of useful methods that can be used to complement a hypothesis-driven analysis (Table 3). Many descriptive methods work best where a small group of taxa is measured only a few times. However, there are examples of descriptive methods being used effectively for larger, more complex examples. Pereira et al. (2017) used a heat map to show compositional changes of 808 bacterial genera, at a single location, sampled 24 times over two years. Certain descriptive methods are useful for only specific types of temporal patterns, e.g., ‘time cores’ is a method that identifies ‘core’ groups of taxa that remain present over time (Vallès et al. 2014). Others, such as Venn diagrams and measures of overlap in composition, are useful only for small numbers of temporal comparisons. We recommend and provide R code (Data S2) for time cores and Venn diagrams because these methods are suitable for assessing hypotheses of taxon replacement while specifically identifying the taxa driving compositional shifts (Tables 2 and 3). Relative abundance distributions (RADs, *a.k.a.*

‘rank abundance curves’), can provide relatively sophisticated visualisations of community dynamics, and rank abundances can be used to quantitatively measure the shift in taxon ranks within samples taken at different time points (Hallett et al. 2016). RADs are plots of the relative abundances (on the y -axis) of all species in a community ranked (on the x -axis) from highest to lowest. They are closely related to k -dominance curves, in which species ranks are shown on the x -axis, but cumulative proportional abundance (dominance) is plotted on the y -axis (Lambhead et al. 1983). If samples taken at different times are plotted separately, the set of RADs can illustrate changes in community structure over time. RADs also can be used to explore community dynamics if a dissimilarity measure, such as Bray-Curtis, is calculated for the species’ ranks at each time point. The resulting temporal dissimilarity of the RADs can be explored by plotting rank dissimilarity against time lag or by ordinating rank dissimilarities in the same way that spatial compositional dissimilarities are examined. For example, temporal changes in RAD dissimilarity can illustrate changes in species evenness over time as might be expected if compositional homogenisation and a decline in evenness occurs because of invasion by an increasingly dominant species (Ruhl 2008).

E. DISCUSSION

There is an ever-expanding set of analyses that use a wide variety of multivariate and univariate methods to quantitatively address temporal questions about communities that continue to be developed (e.g., De Cáceres et al. 2019). One of the key reasons for the proliferation of methods appears to be variability in the length of the time series. For example, if the time series is long enough, methods such as temporal MEMs could be used; however, consistent with other reviews (e.g., Vaughn and Young 2010) and databases (Dornelas et al. 2018), most datasets captured by our literature had relatively few temporal measurements (mean = 14.9, median = 7, mode = 4), requiring very different methods. In other cases, the ability to implement some methods in temporal community ecology still

requires development, such as the application of time-ordered and time-aggregated network analysis for spatiotemporal community datasets (Blonder et al. 2012).

Some analysis types are well established for spatial data, but can be applied easily to temporal data, potentially yielding useful and comparable results for explicit temporal analyses. Further research is required to generate standard extensions and approaches using such methods for temporal data. For example, nestedness (*sensu* Patterson and Atmar 1986) is usually applied to spatial data and the analysis is repeated for different time points; however, there is no reason that we cannot apply this method across times to look at whether sites are becoming more or less nested over time. For instance, if younger sites were nested within older sites, this may indicate that sites were becoming more homogeneous over time.

Our understanding of temporal dynamics in biological communities, and the methods we use to investigate them, have improved greatly over the past 25 years, probably because of increases in computational power and in dataset size. We can continue to improve our understanding by moving away from merely describing community changes, to linking these changes to understanding and testing the processes driving community change and biodiversity; that is, moving from hypothesis-generating methods towards hypothesis-testing methods enabled by well-defined research questions with specific, testable hypotheses that result in clear predictions for observed temporal patterns (e.g., Figs 4-7). Even for exploratory studies of compositional variation through time, we emphasise the benefits of making *a priori* predictions for how we think composition should change through time.

Interactions between community composition and environmental or other factors across space and time are complicated. For example, Brown (2003) showed that temporal variability in species composition of stream invertebrates was reduced by high spatial heterogeneity, which may be because there are increased niches or refugia in changing environmental conditions. Manipulative experiments are useful, but not always possible,

especially for long-lived taxa. Therefore, combining current knowledge of intrinsic and extrinsic factors controlling community change into null models and simulations may be the way forward (Brown and Lawson 2010). Carefully considering sample sizes and appropriate replication of particular factors are also important if we are to properly test hypotheses of species coexistence over time (Philippi et al. 1998). There is a need for us to make useful predictions about how communities may continue to respond to ecosystem disturbances, such as climate change, invasive species, and habitat fragmentation (Hobbs et al. 2007, Kampichler et al. 2012). Many methods commonly used to analyse temporal community dynamics are not quantitative, meaning conclusions can only be narrative descriptions, rather than clear tests of hypotheses. Many of the methods presented here push us to quantitatively assess dynamics, which naturally leads to clearer hypothesis testing.

There is a wide variety of methods available for analysing temporal dynamics of biological communities; however, many studies in our literature review using these methods are descriptive, showing that a large body of literature in community ecology is not focussed on testing quantitative hypotheses of change. Surprisingly few studies in our review presented hypotheses that were more complex than simply predicting that composition would have changed in some way over time, instead of specifying predictions of quantitative and directional patterns of change in community composition. Even additional searching beyond our original search terms that included more specific terminology for compositional change (succession, turnover, synchron*, resilience, homogenization, divergence, convergence) revealed few empirical papers that tested specific, quantitative hypotheses. This search resulted in an additional 1,514 papers of which we read 10% (151 papers). Of these, 60 papers met the inclusion criteria of our original review and only 13 papers tested hypotheses that were more sophisticated than the simple prediction that composition would have changed

in some way over time, e.g., directional change or that the abundances of some taxa would change more than others.

This lack of specific, quantitative hypothesis testing is not surprising because to predict change in composition in a quantitative way, we require an in-depth understanding of a system and how different components of the diversity differ in their ecology or other factors that could cause directional, cyclical, or other differential change in time. In many cases, particularly in the observational, rather than experimental, community ecology literature, this is not the situation. This means that, although such studies may state hypotheses, these hypotheses are not very specific or quantitative. In most cases, they predict some change in response to time or another factor that is correlated with time (such as disturbance or pollution), but the types of temporal change (divergence, convergence, synchrony, etc.) are not specified. We strongly recommend that researchers generate *a priori* hypotheses of compositional change (e.g., Figs. 4-7; Bagchi et al. 2017). These predictions should be used to select appropriate quantitative data analysis methods (Fig. 17). The comprehensive description and hypothesis testing of temporal community patterns requires the use of multiple data analysis methods, but by using standard exploratory data analysis methods followed by testing *a priori* hypotheses, the methods selected can be explicitly focussed on the most appropriate tests for the situation. We hope that this synthesis will allow researchers to make informed decisions about their choice of data analysis for their research question and pave the way for more uniformity in reporting of these analyses and/or innovation in comparability. Our understanding of patterns and drivers of changes in species communities would benefit by having more longer-term and fine-scale studies.

ACKNOWLEDGEMENTS

We thank Audrey Lustig for initial database and R code development. This paper is a contribution from the Harvard Forest LTER program. We thank members of the Forest

1178 Ecology Research Group at Wilfrid Laurier University, Glenda Mendieta-Leiva, Brian
 1179 Inouye, Miguel De Cáceres, and an anonymous reviewer for comments on the manuscript.
 1180 We thank the authors of Porensky et al. (2016a, 2016b) and Gilbert et al. (2012) for making
 1181 their data publicly available. The BCI forest dynamics research project was founded by S.P.
 1182 Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the
 1183 Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama.
 1184 Numerous organizations have provided funding, principally the U.S. National Science
 1185 Foundation, and hundreds of field workers have contributed.

1186 **FUNDING DISCLOSURE**

1187 H. L. Buckley and B. S. Case were supported by a Charles Bullard Fellowship at Harvard
 1188 Forest. A. M. Ellison's work on this project was supported by the Harvard Forest.

1189 **SUPPORTING INFORMATION**

1190 Data S1. Description of attributes, and associated data, extracted from papers in the literature
 1191 review.

1192 Data S2. Example R code for plant community dataset

1193 Data S3. Example R code for BCI dataset

1194 Data S4. Example R code for marine bacterial community dataset (from high throughput
 1195 sequencing)

1196 Appendix S1. Supplementary figures from the literature review.

1197 Appendix S2. Descriptions of methods encountered in the literature review that are not
 1198 included in the main text.

1199

1200 **REFERENCES CITED**

1201 Aguirre, L. F., L. Lens, R. van Damme, and E. Matthysen. 2003. Consistency and variation in
1202 the bat assemblages inhabiting two forest islands within a neotropical savanna in
1203 Bolivia. *Journal of Tropical Ecology* 19:367–374.

1204 Amend, A. S., K. A. Seifert, and T. D. Bruns. 2010. Quantifying microbial communities with
1205 454 pyrosequencing: does read abundance count? *Molecular Ecology* 19:5555–5565.

1206 Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance.
1207 *Austral Ecology* 26:32–46.

1208 Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J.
1209 Sanders, H. V. Cornell, L. S. Comita, K. F. Davies, S. P. Harrison, N. J. B. Kraft, J. C.
1210 Stegen, and N. G. Swenson. 2011. Navigating the multiple meanings of β diversity: A
1211 roadmap for the practicing ecologist. *Ecology Letters* 14:19–28.

1212 Anderson, M. J., and T. J. Willis. 2003. Canonical analysis of principal coordinates: A useful
1213 method of constrained ordination for ecology. *Ecology* 84:511–525.

1214 Angeler, D. G., and R. K. Johnson. 2012. Patterns of temporal community turnover are
1215 spatially synchronous across boreal lakes: Temporal coherence in invertebrate
1216 community turnover. *Freshwater Biology* 57:1782–1793.

1217 Angeler, D. G., O. Viedma, and J. M. Moreno. 2009. Statistical performance and information
1218 content of time lag analysis and redundancy analysis in time series modeling. *Ecology*
1219 90:3245–3257.

1220 Auber, A., M. Travers-Trolet, M. C. Villanueva, and B. Ernande. 2017. A new application of
1221 principal response curves for summarizing abrupt and cyclic shifts of communities
1222 over space. *Ecosphere* 8:e02023.

1223 Azeria, E. T., A. Carlson, T. Pärt, and C. G. Wiklund. 2006. Temporal dynamics and
1224 nestedness of an oceanic island bird fauna. *Global Ecology and Biogeography*
1225 15:328–338.

1226 Azeria, E. T., and J. Kolasa. 2008. Nestedness, niche metrics and temporal dynamics of a
1227 metacommunity in a dynamic natural model system. *Oikos* 117:1006–1019.

1228 Bagchi, S., N. J. Singh, D. D. Briske, B. T. Bestelmeyer, M. P. McClaran, and K. Murthy.
1229 2017. Quantifying long-term plant community dynamics with movement models:
1230 implications for ecological resilience. *Ecological Applications* 27:1514–1528.

1231 Bar-Massada, A., and L. Hadar. 2017. Grazing and temporal turnover in herbaceous
1232 communities in a Mediterranean landscape. *Journal of Vegetation Science* 28:270–
1233 280.

1234 Barros, F., H. Blanchet, K. Hammerstrom, P.-G. Sauriau, and J. Oliver. 2014. A framework
1235 for investigating general patterns of benthic β -diversity along estuaries. *Estuarine,
1236 Coastal and Shelf Science* 149:223–231.

1237 Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity:
1238 Partitioning beta diversity. *Global Ecology and Biogeography* 19:134–143.

1239 Baselga, A. 2012. The relationship between species replacement, dissimilarity derived from
1240 nestedness, and nestedness: Species replacement and nestedness. *Global Ecology and
1241 Biogeography* 21:1223–1232.

1242 Baselga, A. 2017. Partitioning abundance-based multiple-site dissimilarity into components:
1243 Balanced variation in abundance and abundance gradients. *Methods in Ecology and
1244 Evolution* 8:799–808.

1245 Baselga, A., and C. D. L. Orme. 2012. betapart : an R package for the study of beta diversity:
1246 Betapart package. *Methods in Ecology and Evolution* 3:808–812.

1247 Bashan, A., T. E. Gibson, J. Friedman, V. J. Carey, S. T. Weiss, E. L. Hohmann, and Y.-Y.
1248 Liu. 2016. Universality of human microbial dynamics. *Nature* 534:259–262.

1249 Bêche, L. A., and V. H. Resh. 2007. Short-term climatic trends affect the temporal variability
1250 of macroinvertebrates in California ‘Mediterranean’ streams. *Freshwater Biology*
1251 52:2317–2339.

1252 Beisner, B. E., A. R. Ives, and S. R. Carpenter. 2003. The effects of an exotic fish invasion on
1253 the prey communities of two lakes. *Journal of Animal Ecology* 72:331–342.

1254 Berthon, V., B. Alric, F. Rimet, and M.-E. Perga. 2014. Sensitivity and responses of diatoms
1255 to climate warming in lakes heavily influenced by humans. *Freshwater Biology*
1256 59:1755–1767.

1257 Bestelmeyer, B. T., A. M. Ellison, W. R. Fraser, K. B. Gorman, S. J. Holbrook, C. M. Laney,
1258 M. D. Ohman, D. P. C. Peters, F. C. Pillsbury, A. Rassweiler, R. J. Schmitt, and S.
1259 Sharma. 2011. Analysis of abrupt transitions in ecological systems. *Ecosphere*
1260 2:art129.

1261 Blonder, B., T. W. Wey, A. Dornhaus, R. James, and A. Sih. 2012. Temporal dynamics and
1262 network analysis. *Methods in Ecology and Evolution* 3:958–972.

1263 Borcard, D., F. Gillet, and P. Legendre. 2011. *Numerical Ecology with R*. Springer, New
1264 York.

1265 Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of
1266 principal coordinates of neighbour matrices. *Ecological Modelling* 153:51–68.

1267 Boulinier, T., J. D. Nichols, J. E. Hines, J. R. Sauer, C. H. Flather, and K. H. Pollock. 2001.
1268 Forest fragmentation and bird community dynamics: Inference at regional scales.
1269 *Ecology* 82:1159–1169.

1270 ter Braak, C. J. F., and P. Šmilauer. 2015. Topics in constrained and unconstrained
1271 ordination. *Plant Ecology* 216:683–696.

1272 Broadway, K. J., M. Pyron, J. R. Gammon, and B. A. Murry. 2015. Shift in a large river fish
1273 assemblage: Body-size and trophic structure dynamics. *PLOS ONE* 10:e0124954.

1274 Brown, B. L. 2003. Spatial heterogeneity reduces temporal variability in stream insect
1275 communities. *Ecology Letters* 6:316–325.

1276 Brown, B. L., and R. L. Lawson. 2010. Habitat heterogeneity and activity of an omnivorous
1277 ecosystem engineer control stream community dynamics. *Ecology* 91:1799–1810.

1278 Caruso, T., V. Trokhymets, R. Bargagli, and P. Convey. 2013. Biotic interactions as a
1279 structuring force in soil communities: Evidence from the micro-arthropods of an
1280 Antarctic moss model system. *Oecologia* 172:495–503.

1281 Chao, A., R. L. Chazdon, R. K. Colwell, and T.-J. Shen. 2006. Abundance-based similarity
1282 indices and their estimation when there are unseen species in samples. *Biometrics*
1283 62:361–371.

1284 Chao, A., N. J. Gotelli, T. C. Hsieh, E. L. Sander, K. H. Ma, R. K. Colwell, and A. M.
1285 Ellison. 2014. Rarefaction and extrapolation with Hill numbers: a framework for
1286 sampling and estimation in species diversity studies. *Ecological Monographs* 84:45–
1287 67.

1288 Chapman, J. I., and R. W. McEwan. 2016. Thirty years of compositional change in an old-
1289 growth temperate forest: The role of topographic gradients in oak-maple dynamics.
1290 *PLOS ONE* 11:e0160238.

1291 Clavero, M., and E. Garcia-Berthou. 2006. Homogenization dynamics and introduction
1292 routes of invasive freshwater fish in the Iberian Peninsula. *Ecological Applications*
1293 16:2313–2324.

1294 Cleland, E. E., S. L. Collins, T. L. Dickson, E. C. Farrer, K. L. Gross, L. A. Gherardi, L. M.
1295 Hallett, R. J. Hobbs, J. S. Hsu, L. Turnbull, and K. N. Suding. 2013. Sensitivity of
1296 grassland plant community composition to spatial vs. temporal variation in
1297 precipitation. *Ecology* 94:1687–1696.

1298 Collins, S. L. 1992. Fire frequency and community heterogeneity in tallgrass prairie
1299 vegetation. *Ecology* 73:2001–2006.

1300 Collins, S. L. 2000. Disturbance frequency and community stability in native tallgrass prairie.
1301 *The American Naturalist* 155:311–325.

1302 Collins, S. L., F. Micheli, and L. Hartt. 2000. A method to determine rates and patterns of
1303 variability in ecological communities. *Oikos* 91:285–293.

1304 Collins, S. L., and M. D. Smith. 2006. Scale-dependent interaction of fire and grazing on
1305 community heterogeneity in tallgrass prairie. *Ecology* 87:2058–2067.

1306 Collins, S. L., and Y. Xia. 2015. Long-term dynamics and hotspots of change in a desert
1307 grassland plant community. *The American Naturalist* 185:E30–E43.

1308 Condit, R. 1998. *Tropical Forest Census Plots*. Springer-Verlag and R. G. Landes Company,
1309 Berlin, Germany, and Georgetown, Texas.

1310 Connell, J. H., and R. O. Slatyer. 1977. Mechanisms of succession in natural communities
1311 and their role in community stability and organization. *The American Naturalist*
1312 111:1119–1144.

1313 Cotton, T. E. A., A. H. Fitter, R. M. Miller, A. J. Dumbrell, and T. Helgason. 2015. Fungi in
1314 the future: interannual variation and effects of atmospheric change on arbuscular
1315 mycorrhizal fungal communities. *New Phytologist* 205:1598–1607.

1316 Day, N. J., and H. L. Buckley. 2013. Twenty-five years of plant community dynamics and
1317 invasion in New Zealand tussock grasslands. *Austral Ecology* 38:688–699.

1318 Day, N. J., S. Carrière, and J. L. Baltzer. 2017. Annual dynamics and resilience in post-fire
1319 boreal understory vascular plant communities. *Forest Ecology and Management*
1320 401:264–272.

1321 De Cáceres, M., L. Coll, P. Legendre, R. B. Allen, S. K. Wiser, M.-J. Fortin, R. Condit, and
1322 S. Hubbell. 2019. Trajectory analysis in community ecology. *Ecological*
1323 *Monographs*:e01350.

1324 De Cáceres, M., X. Font, and F. Oliva. 2010. The management of vegetation classifications
1325 with fuzzy clustering: Fuzzy clustering in vegetation classifications. *Journal of*
1326 *Vegetation Science* 21:1138–1151.

1327 Diamond, J. M. 1969. Avifaunal equilibria and species turnover rates on the Channel Islands
1328 of California. *Proceedings of the National Academy of Sciences* 64:57–63.

1329 Dimitriu, P. A., G. Boyce, A. Samarakoon, M. Hartmann, P. Johnson, and W. W. Mohn.
1330 2013. Temporal stability of the mouse gut microbiota in relation to innate and
1331 adaptive immunity: Mouse gut microbiota dynamics. *Environmental Microbiology*
1332 *Reports* 5:200–210.

1333 Dorazio, R. M., E. F. Connor, and R. A. Askins. 2015. Estimating the Effects of Habitat and
1334 Biological Interactions in an Avian Community. *PLOS ONE* 10:e0135987.

1335 Dornelas, M., L. H. Antão, F. Moyes, A. E. Bates, A. E. Magurran, D. Adam, A. A.
1336 Akhmetzhanova, W. Appeltans, J. M. Arcos, H. Arnold, N. Ayyappan, G. Badihi, A.
1337 H. Baird, M. Barbosa, T. E. Barreto, C. Bässler, A. Bellgrove, J. Belmaker, L.
1338 Benedetti-Cecchi, B. J. Bett, A. D. Bjorkman, M. Błażewicz, S. A. Blowes, C. P.
1339 Bloch, T. C. Bonebrake, S. Boyd, M. Bradford, A. J. Brooks, J. H. Brown, H.
1340 Bruelheide, P. Budy, F. Carvalho, E. Castañeda-Moya, C. A. Chen, J. F. Chamblee, T.
1341 J. Chase, L. Siegwart Collier, S. K. Collinge, R. Condit, E. J. Cooper, J. H. C.
1342 Cornelissen, U. Cotano, S. Kyle Crow, G. Damasceno, C. H. Davies, R. A. Davis, F.
1343 P. Day, S. Degraer, T. S. Doherty, T. E. Dunn, G. Durigan, J. E. Duffy, D. Edelist, G.
1344 J. Edgar, R. Elahi, S. C. Elmendorf, A. Enemar, S. K. M. Ernest, R. Escibano, M.
1345 Estiarte, B. S. Evans, T.-Y. Fan, F. Turini Farah, L. Loureiro Fernandes, F. Z.

1346 Farneda, A. Fidelis, R. Fitt, A. M. Fosaa, G. A. Daher Correa Franco, G. E. Frank, W.
1347 R. Fraser, H. García, R. Cazzolla Gatti, O. Givan, E. Gorgone-Barbosa, W. A. Gould,
1348 C. Gries, G. D. Grossman, J. R. Gutierrez, S. Hale, M. E. Harmon, J. Harte, G.
1349 Haskins, D. L. Henshaw, L. Hermanutz, P. Hidalgo, P. Higuchi, A. Hoey, G. Van
1350 Hoey, A. Hofgaard, K. Holeck, R. D. Hollister, R. Holmes, M. Hoogenboom, C.
1351 Hsieh, S. P. Hubbell, F. Huettmann, C. L. Huffard, A. H. Hurlbert, N. Macedo
1352 Ivanauskas, D. Janík, U. Jandt, A. Jazdzewska, T. Johannessen, J. Johnstone, J. Jones,
1353 F. A. M. Jones, J. Kang, T. Kartawijaya, E. C. Keeley, D. A. Kelt, R. Kinnear, K.
1354 Klanderud, H. Knutsen, C. C. Koenig, A. R. Kortz, K. Král, L. A. Kuhnz, C.-Y. Kuo,
1355 D. J. Kushner, C. Laguionie-Marchais, L. T. Lancaster, C. Min Lee, J. S. Lefcheck, E.
1356 Lévesque, D. Lightfoot, F. Lloret, J. D. Lloyd, A. López-Baucells, M. Louzao, J. S.
1357 Madin, B. Magnússon, S. Malamud, I. Matthews, K. P. McFarland, B. McGill, D.
1358 McKnight, W. O. McLarney, J. Meador, P. L. Meserve, D. J. Metcalfe, C. F. J.
1359 Meyer, A. Michelsen, N. Milchakova, T. Moens, E. Moland, J. Moore, C. Mathias
1360 Moreira, J. Müller, G. Murphy, I. H. Myers-Smith, R. W. Myster, A. Naumov, F.
1361 Neat, J. A. Nelson, M. Paul Nelson, S. F. Newton, N. Norden, J. C. Oliver, E. M.
1362 Olsen, V. G. Onipchenko, K. Pabis, R. J. Pabst, A. Paquette, S. Pardede, D. M.
1363 Paterson, R. Péliissier, J. Peñuelas, A. Pérez-Matus, O. Pizarro, F. Pomati, E. Post, H.
1364 H. T. Prins, J. C. Priscu, P. Provoost, K. L. Prudic, E. Pulliainen, B. R. Ramesh, O.
1365 Mendivil Ramos, A. Rassweiler, J. E. Rebelo, D. C. Reed, P. B. Reich, S. M.
1366 Remillard, A. J. Richardson, J. P. Richardson, I. van Rijn, R. Rocha, V. H. Rivera-
1367 Monroy, C. Rixen, K. P. Robinson, R. Ribeiro Rodrigues, D. de Cerqueira Rossa-
1368 Feres, L. Rudstam, H. Ruhl, C. S. Ruz, E. M. Sampaio, N. Rybicki, A. Rypel, S. Sal,
1369 B. Salgado, F. A. M. Santos, A. P. Savassi-Coutinho, S. Scanga, J. Schmidt, R.
1370 Schooley, F. Setiawan, K.-T. Shao, G. R. Shaver, S. Sherman, T. W. Sherry, J.

1371 Siciński, C. Sievers, A. C. da Silva, F. Rodrigues da Silva, F. L. Silveira, J. Slingsby,
1372 T. Smart, S. J. Snell, N. A. Soudzilovskaia, G. B. G. Souza, F. Maluf Souza, V.
1373 Castro Souza, C. D. Stallings, R. Stanforth, E. H. Stanley, J. Mauro Sterza, M.
1374 Stevens, R. Stuart-Smith, Y. Rondon Suarez, S. Supp, J. Yoshio Tamashiro, S.
1375 Tarigan, G. P. Thiede, S. Thorn, A. Tolvanen, M. Teresa Zugliani Toniato, Ø.
1376 Totland, R. R. Twilley, G. Vaitkus, N. Valdivia, M. I. Vallejo, T. J. Valone, C. Van
1377 Colen, J. Vanaverbeke, F. Venturoli, H. M. Verheye, M. Vianna, R. P. Vieira, T.
1378 Vrška, C. Quang Vu, L. Van Vu, R. B. Waide, C. Waldock, D. Watts, S. Webb, T.
1379 Wesolowski, E. P. White, C. E. Widdicombe, D. Wilgers, R. Williams, S. B.
1380 Williams, M. Williamson, M. R. Willig, T. J. Willis, S. Wipf, K. D. Woods, E. J.
1381 Woehler, K. Zawada, and M. L. Zettler. 2018. BioTIME: A database of biodiversity
1382 time series for the Anthropocene. *Global Ecology and Biogeography* 27:760–786.
1383 Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive
1384 framework for principal coordinate analysis of neighbour matrices (PCNM).
1385 *Ecological Modelling* 196:483–493.
1386 Ellison, A. 2010. Partitioning diversity. *Ecology* 91:1962–1963.
1387 Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalized dissimilarity
1388 modelling to analyse and predict patterns of beta diversity in regional biodiversity
1389 assessment. *Diversity and Distributions* 13:252–264.
1390 Franca, A., S. Caredda, F. Sanna, F. Fava, and G. Seddaiu. 2016. Early plant community
1391 dynamics following overseeding for the rehabilitation of a Mediterranean
1392 silvopastoral system. *Grassland Science* 62:81–91.
1393 Gaston, K. J., and B. H. McArdle. 1994. The temporal variability of animal abundances:
1394 measures, methods and patterns. *Philosophical Transactions of the Royal Society B:*
1395 *Biological Sciences* 345:335–358.

1396 Gilbert, J. A., J. A. Steele, J. G. Caporaso, L. Steinbrück, J. Reeder, B. Temperton, S. Huse,
1397 A. C. McHardy, R. Knight, and I. Joint. 2012. Defining seasonal marine microbial
1398 community dynamics. *The ISME journal* 6:298.

1399 Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81:2606–
1400 2621.

1401 Gotelli, N. J. 2001. Research frontiers in null model analysis. *Global Ecology and*
1402 *Biogeography* 10:337–343.

1403 Gotelli, N. J., and G. L. Entsminger. 2001. Swap and fill algorithms in null model analysis:
1404 rethinking the knight’s tour. *Oecologia* 129:281–291.

1405 Gotelli, N. J., and G. R. Graves. 1996. *Null Models in Ecology*. Smithsonian Institution
1406 Press, Washington and London.

1407 Gouhier, T. C., and F. Guichard. 2014. Synchrony: Quantifying variability in space and time.
1408 *Methods in Ecology and Evolution* 5:524–533.

1409 Hallett, L. M., S. K. Jones, A. A. M. MacDonald, M. B. Jones, D. F. B. Flynn, J. Ripplinger,
1410 P. Slaughter, C. Gries, and S. L. Collins. 2016. codyn: An R package of community
1411 dynamics metrics. *Methods in Ecology and Evolution* 7:1146–1151.

1412 Hampton, S. E., E. E. Holmes, L. P. Scheef, M. D. Scheuerell, S. L. Katz, D. E. Pendleton,
1413 and E. J. Ward. 2013. Quantifying effects of abiotic and biotic drivers on community
1414 dynamics with multivariate autoregressive (MAR) models. *Ecology* 94:2663–2669.

1415 Hampton, S. E., L. R. Izmet’Eva, M. V. Moore, S. L. Katz, B. Dennis, and E. A. Silow.
1416 2008. Sixty years of environmental change in the world’s largest freshwater lake -
1417 Lake Baikal, Siberia. *Global Change Biology* 14:1947–1958.

1418 Hobbs, R. J., S. Yates, and H. A. Mooney. 2007. Long-term data reveal complex dynamics in
1419 grassland in relation to climate and disturbance. *Ecological Monographs* 77:545–568.

1420 Holmes, E. E., E. J. Ward, and K. Wills. 2012. MARSS: Multivariate Autoregressive State-
1421 space models for analyzing time-series data. *R Journal* 4:11–19.

1422 Houlahan, J. E., D. J. Currie, K. Cottenie, G. S. Cumming, S. K. M. Ernest, C. S. Findlay, S.
1423 D. Fuhlendorf, U. Gaedke, P. Legendre, J. J. Magnuson, B. H. McArdle, E. H.
1424 Muldavin, D. Noble, R. Russell, R. D. Stevens, T. J. Willis, I. P. Woiwod, and S. M.
1425 Wondzell. 2007. Compensatory dynamics are rare in natural ecological communities.
1426 *Proceedings of the National Academy of Sciences* 104:3273–3277.

1427 Hubbell, S. P. 1979. Tree dispersion, abundance, and diversity in a tropical dry forest.
1428 *Science* 203:1299–1309.

1429 Hubbell, S. P., R. Condit, and R. B. Foster. 2005. Barro Colorado Forest Census Plot Data.

1430 Hubbell, S. P., R. B. Foster, S. T. O’Brien, K. E. Harms, R. Condit, B. Wechsler, S. J.
1431 Wright, and S. Loo de Lao. 1999. Light-gap disturbances, recruitment limitation, and
1432 tree diversity in a neotropical forest. *Science* 283:554–557.

1433 Hui, C., and M. A. McGeoch. 2014. Zeta diversity as a concept and metric that unifies
1434 incidence-based biodiversity patterns. *The American Naturalist* 184:684–694.

1435 Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments.
1436 *Ecological Monographs* 54:187–211.

1437 Ives, A. R., B. Dennis, K. L. Cottingham, and S. R. Carpenter. 2003. Estimating community
1438 stability and ecological interactions from time-series data. *Ecological monographs*
1439 73:301–330.

1440 Jackson, D. A. 1995. PROTEST: A PROcrustean Randomization TEST of community
1441 environment concordance. *Écoscience* 2:297–303.

1442 Johnstone, I. M., and D. M. Titterton. 2009. Statistical challenges of high-dimensional
1443 data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and*
1444 *Engineering Sciences* 367:4237–4253.

1445 Kalyuzhny, M., and N. M. Shnerb. 2017. Dissimilarity-overlap analysis of community
1446 dynamics: Opportunities and pitfalls. *Methods in Ecology and Evolution* 8:1764–
1447 1773.

1448 Kampichler, C., D. G. Angeler, R. T. Holmes, A. Leito, S. Svensson, H. P. van der Jeugd,
1449 and T. Wesołowski. 2014. Temporal dynamics of bird community composition: An
1450 analysis of baseline conditions from long-term data. *Oecologia* 175:1301–1313.

1451 Kampichler, C., C. A. M. van Turnhout, V. Devictor, and H. P. van der Jeugd. 2012. Large-
1452 scale changes in community composition: Determining land use and climate change
1453 signals. *PLoS ONE* 7:e35272.

1454 Kara, E., and A. Shade. 2009. Temporal dynamics of South End Tidal Creek (Sapelo Island,
1455 Georgia) bacterial communities. *Applied and Environmental Microbiology* 75:1058–
1456 1064.

1457 Kardol, P., D. E. Todd, P. J. Hanson, and P. J. Mulholland. 2010. Long-term successional
1458 forest dynamics: species and community responses to climatic variability. *Journal of*
1459 *Vegetation Science*.

1460 Kent, M. 2012. *Vegetation Description and Data Analysis: A practical approach*. Second
1461 Edition. Wiley-Blackwell, Chichester, UK.

1462 Kidé, S. O., C. Manté, L. Dubroca, H. Demarcq, and B. Mérigot. 2015. Spatio-temporal
1463 dynamics of exploited groundfish species assemblages faced to environmental and
1464 fishing forcings: insights from the Mauritanian exclusive economic zone. *PloS one*
1465 10:e0141566.

1466 Lamshead, P. J. D., H. M. Platt, and K. M. Shaw. 1983. The detection of differences among
1467 assemblages of marine benthic species based on an assessment of dominance and
1468 diversity. *Journal of Natural History* 17:859–874.

- 1469 Lamy, T., P. Legendre, Y. Chancerelle, G. Siu, and J. Claudet. 2015. Understanding the
1470 spatio-temporal response of coral reef fish communities to natural disturbances:
1471 Insights from beta-diversity decomposition. *PLOS ONE* 10:e0138696.
- 1472 Latombe, G., C. Hui, and M. A. McGeoch. 2017a. Multi-site generalised dissimilarity
1473 modelling: using zeta diversity to differentiate drivers of turnover in rare and
1474 widespread species. *Methods in Ecology and Evolution* 8:431–442.
- 1475 Latombe, G., M. A. McGeoch, D. A. Nipperess, and C. Hui. 2017b. zetadiv: Functions to
1476 Compute Compositional Turnover Using Zeta Diversity. R package version 1.0.
1477 <https://CRAN.R-project.org/package=zetadiv>.
- 1478 Legendre, P., and M. De Cáceres. 2013. Beta diversity as the variance of community data:
1479 dissimilarity coefficients and partitioning. *Ecology Letters* 16:951–963.
- 1480 Legendre, P., and E. Gallagher. 2001. Ecologically meaningful transformations for ordination
1481 of species data. *Oecologia* 129:271–280.
- 1482 Legendre, P., and O. Gauthier. 2014. Statistical methods for temporal and space-time analysis
1483 of community composition data. *Proceedings of the Royal Society B: Biological*
1484 *Sciences* 281:20132728–20132728.
- 1485 Legendre, P., and L. F. J. Legendre. 2012. *Numerical Ecology*. Third Edition. Elsevier,
1486 Radarweg, The Netherlands.
- 1487 Lellouch, L., S. Pavoine, F. Jiguet, H. Glotin, and J. Sueur. 2014. Monitoring temporal
1488 change of bird communities with dissimilarity acoustic indices. *Methods in Ecology*
1489 *and Evolution* 5:495–505.
- 1490 Lepš, J., and P. Šmilauer. 2003. *Multivariate Analysis of Ecological Data Using CANOCO*.
1491 Cambridge University Press, Cambridge.
- 1492 Lewis, S. L., and M. A. Maslin. 2015. Defining the Anthropocene. *Nature* 519:171–180.

1493 Loreau, M., and C. de Mazancourt. 2008. Species synchrony and its drivers: neutral and
1494 nonneutral community dynamics in fluctuating environments. *The American*
1495 *Naturalist* 172:E48–E66.

1496 Lozupone, C. A., M. Hamady, S. T. Kelley, and R. Knight. 2007. Quantitative and qualitative
1497 diversity measures lead to different insights into factors that structure microbial
1498 communities. *Applied and Environmental Microbiology* 73:1576–1585.

1499 Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing
1500 microbial communities. *Applied and Environmental Microbiology* 71:8228–8235.

1501 Magalhães, M. F., P. Beja, I. J. Schlosser, and M. J. Collares-Pereira. 2007. Effects of multi-
1502 year droughts on fish assemblages of seasonally drying Mediterranean streams.
1503 *Freshwater Biology* 52:1494–1510.

1504 Magurran, A. E. 2004. *Measuring Biological Diversity*. Blackwell, Malden.

1505 Magurran, A. E. 2013. *Measuring Biological Diversity*. John Wiley and Sons, Oxford.

1506 McCune, B., and J. B. Grace. 2002. *Analysis of Ecological Communities*. MjM Software
1507 Design, Gleneden Beach.

1508 McGeoch, M. A., G. Latombe, N. R. Andrew, S. Nakagawa, D. A. Nipperess, M. Roige, E. .
1509 Marzinelli, A. H. Campbell, A. Vergés, T. Thomas, P. D. Steinberg, K. E. Selwood,
1510 and C. Hui. 2017. The application of zeta diversity as a continuous measure of
1511 compositional change in ecology. *bioRxiv* 216580.

1512 Micheli, F., K. L. Cottingham, J. Bascompte, O. N. Bjornstad, G. L. Eckert, J. M. Fischer, T.
1513 H. Keitt, B. E. Kendall, J. L. Klug, and J. A. Rusak. 1999. The dual nature of
1514 community variability. *Oikos* 85:161–169.

1515 Mutshinda, C. M., R. B. O’Hara, and I. P. Woiwod. 2009. What drives community
1516 dynamics? *Proceedings of the Royal Society B: Biological Sciences* 276:2923–2929.

1517 Noll, M., D. Matthies, P. Frenzel, M. Derakshani, and W. Liesack. 2005. Succession of
1518 bacterial community structure and diversity in a paddy soil oxygen gradient.
1519 *Environmental Microbiology* 7:382–395.

1520 Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin,
1521 R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H.
1522 Wagner. 2017. *vegan: Community Ecology Package*. R package version 2.4-2.
1523 <http://CRAN.R-project.org/package=vegan>.

1524 O’Sullivan, D. J., P. D. Cotter, O. O’Sullivan, L. Giblin, P. L. H. McSweeney, and J. J.
1525 Sheehan. 2015. Temporal and spatial differences in microbial composition during the
1526 manufacture of a Continental-type cheese. *Applied and Environmental Microbiology*
1527 81:2525–2533.

1528 Ovaskainen, O., G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther, and N. Abrego.
1529 2017. How are species interactions structured in species-rich communities? A new
1530 method for analysing time-series data. *Proceedings of the Royal Society B: Biological*
1531 *Sciences* 284:20170768.

1532 Palmer, M. 2019. *Ordination Methods for Ecologists*. <http://ordination.okstate.edu/>.

1533 Patterson, B. D., and W. Atmar. 1986. Nested subsets and the structure of insular mammalian
1534 faunas and archipelagos. *Biological journal of the Linnean society* 28:65–82.

1535 Pedersen, E. J., P. L. Thompson, R. A. Ball, M.-J. Fortin, T. C. Gouhier, H. Link, C. Moritz,
1536 H. Nenzen, R. R. E. Stanley, Z. E. Taranu, A. Gonzalez, F. Guichard, and P. Pepin.
1537 2017. Signatures of the collapse and incipient recovery of an overexploited marine
1538 ecosystem. *Royal Society Open Science* 4:170215.

1539 Penny, W., and L. Harrison. 2006. Multivariate autoregressive models. *Statistical Parametric*
1540 *Mapping: The Analysis of Functional Brain Images*, eds K. J. Friston, JT Ashburner,
1541 SJ Kiebel, TE Nichols, and WD Penny (London: Elsevier):534–540.

- 1542 Pereira, R. P. A., J. Peplies, M. G. Höfle, and I. Brettar. 2017. Bacterial community dynamics
1543 in a cooling tower with emphasis on pathogenic bacteria and *Legionella* species using
1544 universal and genus-specific deep sequencing. *Water Research* 122:363–376.
- 1545 Peres-Neto, P. R., and D. A. Jackson. 2001. How well do multivariate data sets match? The
1546 advantages of a Procrustean superimposition approach over the Mantel test.
1547 *Oecologia* 129:169–178.
- 1548 Philippi, T. E., P. M. Dixon, and B. E. Taylor. 1998. Detecting trends in species composition.
1549 *Ecological Applications* 8:300–308.
- 1550 Podani, J., and D. Schmera. 2011. A new conceptual and methodological framework for
1551 exploring and explaining pattern in presence - absence data. *Oikos* 120:1625–1638.
- 1552 Porensky, L. M., K. E. Mueller, D. J. Augustine, and J. D. Derner. 2016a. Thresholds and
1553 gradients in a semi-arid grassland: long-term grazing treatments induce slow,
1554 continuous and reversible vegetation change. *Journal of Applied Ecology* 53:1013–
1555 1022.
- 1556 Porensky, L. M., K. E. Mueller, D. J. Augustine, and J. D. Derner. 2016b. Data from:
1557 Thresholds and gradients in a semi-arid grassland: long-term grazing treatments
1558 induce slow, continuous and reversible vegetation change.
1559 <https://doi.org/10.5061/dryad.3m8p1>.
- 1560 Quinn, M. A., S. E. Halbert, and L. Williams. 1991. Spatial and temporal changes in aphid
1561 (Homoptera: Aphididae) species assemblages collected with suction traps in Idaho.
1562 *Journal of Economic Entomology* 84:1710–1716.
- 1563 R Core Development Team. 2017. R: A language and environment for statistical computing
1564 v. 3.3.3. R Foundation for Statistical Computing, Vienna.
- 1565 Rayner, T. S., B. J. Pusey, and R. G. Pearson. 2009. Spatio-temporal dynamics of fish
1566 feeding in the lower Mulgrave River, north-eastern Queensland: The influence of

1567 seasonal flooding, instream productivity and invertebrate abundance. *Marine and*
1568 *Freshwater Research* 60:97.

1569 Rojas-Sandoval, J., E. Meléndez-Ackerman, and D. S. Fernández. 2012. Plant community
1570 dynamics of a tropical semi-arid system following experimental removals of an exotic
1571 grass. *Applied Vegetation Science* 15:513–524.

1572 Ruhl, H. A. 2008. Community change in the variable resource habitat of the abyssal
1573 Northeast Pacific. *Ecology* 89:991–1000.

1574 Rusak, J. A., N. D. Yan, K. M. Somers, and D. J. McQueen. 1999. The temporal coherence of
1575 zooplankton population abundances in neighboring north-temperate lakes. *The*
1576 *American Naturalist* 153:46–58.

1577 Schaefer, J., K. Gido, and M. Smith. 2005. A test for community change using a null model
1578 approach. *Ecological Applications* 15:1761–1771.

1579 Shenhav, L., O. Furman, I. Mizrahi, and E. Halperin. 2017. Modeling the temporal dynamics
1580 of the gut microbial community in adults and infants. *bioRxiv*:212993.

1581 Shertzer, K. W., E. H. Williams, and J. C. Taylor. 2009. Spatial structure and temporal
1582 patterns in a large marine ecosystem: Exploited reef fishes of the southeast United
1583 States. *Fisheries Research* 100:126–133.

1584 Shi, S., D. Zou, Q. Wang, X. Xia, T. Zheng, C. Wu, and M. Gao. 2016. Responses of
1585 ammonia-oxidizing bacteria community composition to temporal changes in
1586 physicochemical parameters during food waste composting. *RSC Adv.* 6:9541–9548.

1587 Shimadzu, H., M. Dornelas, and A. E. Magurran. 2015. Measuring temporal turnover in
1588 ecological communities. *Methods in Ecology and Evolution* 6:1384–1394.

1589 Stead, T. K., J. M. Schmid-Araya, and A. G. Hildrew. 2003. All creatures great and small:
1590 patterns in the stream benthos across a wide range of metazoan body size. *Freshwater*
1591 *Biology* 48:532–547.

1592 Sugiura, S., T. Tsuru, and Y. Yamaura. 2013. Effects of an invasive alien tree on the diversity
1593 and temporal dynamics of an insect assemblage on an oceanic island. *Biological*
1594 *Invasions* 15:157–169.

1595 Thibaut, L. M., S. R. Connolly, and H. P. A. Sweatman. 2012. Diversity and stability of
1596 herbivorous fishes on coral reefs. *Ecology* 93:891–901.

1597 Thrush, S. F., J. Halliday, J. E. Hewitt, and A. M. Lohrer. 2008. The effects of habitat loss,
1598 fragmentation, and community homogenization on resilience in estuaries. *Ecological*
1599 *Applications* 18:12–21.

1600 Tilman, D. 1999. The ecological consequences of changes in biodiversity: A search for
1601 general principles. *Ecology* 80:1455–1474.

1602 Tworek, S. 2004. Factors Affecting Temporal Dynamics of Avian Assemblages in a
1603 Heterogeneous Landscape. *Acta Ornithologica* 39:155–163.

1604 Ulrich, W., and N. J. Gotelli. 2007. Null model analysis of species nestedness patterns.
1605 *Ecology* 88:1824–1831.

1606 Ulrich, W., and N. J. Gotelli. 2010. Null model analysis of species associations using
1607 abundance data. *Ecology* 91:3384–3397.

1608 Vallès, Y., A. Artacho, A. Pascual-García, M. L. Ferrús, M. J. Gosálbes, J. J. Abellán, and M.
1609 P. Francino. 2014. Microbial succession in the gut: Directional trends of taxonomic
1610 and functional change in a birth cohort of Spanish infants. *PLoS Genetics*
1611 10:e1004406.

1612 Van den Brink, P. J., and C. J. Ter Braak. 1999. Principal response curves: Analysis of time-
1613 dependent multivariate responses of biological community to stress. *Environmental*
1614 *Toxicology and Chemistry* 18:138–148.

1615 Vaughn, K. J., and T. P. Young. 2010. Contingent conclusions: Year of initiation influences
1616 ecological field experiments, but temporal replication is rare. *Restoration Ecology*
1617 18:59–64.

1618 Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California.
1619 *Ecological Monographs* 30:279.

1620 Xiong, J., J. Zhu, K. Wang, X. Wang, X. Ye, L. Liu, Q. Zhao, M. Hou, L. Qiuqian, and D.
1621 Zhang. 2014. The temporal scaling of bacterioplankton composition: high turnover
1622 and predictability during shrimp cultivation. *Microbial Ecology* 67:256–264.

1623 Yamamura, K., M. Yokozawa, M. Nishimori, Y. Ueda, and T. Yokosuka. 2006. How to
1624 analyze long-term insect population dynamics under climate change: 50-year data of
1625 three insect pests in paddy fields. *Population Ecology* 48:31–48.

1626 Zar, J. H. 1996. *Biostatistical Analysis*. Third. Prentice Hall, New Jersey.

1627 Zuur, A. F., E. N. Ieno, and C. S. Elphick. 2010. A protocol for data exploration to avoid
1628 common statistical problems: Data exploration. *Methods in Ecology and Evolution*
1629 1:3–14.

1630

1631

1632 **TABLES**

1633

1634 **Table 1:** Recommended steps for exploratory data analysis (EDA) of community data prior to selecting analysis methods for temporal data. R
 1635 code for the methods is provided in the supplementary materials Data S2-4.

Step	Description	Rationale	Methods
1	Count the number of taxa, the number of samples and the number of empty samples	Some ordination methods are influenced by rare taxa; identify empty samples; determine replication for significance testing against explanatory variables	Row and column totals of the sample $\times$ taxon matrix. Plot the relative abundance distributions.
2	Examine taxon frequency distributions and response curves	Some ordination methods assume either unimodal or linear pattern of frequency distribution	Plot the abundance of species across time or other explanatory variables.

3	Determine the ratio of temporal to spatial variation in composition	If temporal change is to be detected within different spatial sampling locations, the analysis must adequately tease apart the variation in time and space. Where little spatial variation exists, spatial samples could be averaged at each time.	<p>Mean spatial and temporal turnover should be calculated following Collins et al. (2000) to allow a ‘global’ comparison of the amount of compositional variability among samples from each time and among times (see Hallett et al. 2016 for details); it can be calculated for different subsets of samples or times for which different predictions about temporal dynamics have been made, e.g., different experimental treatments.</p> <p>Ordination approaches, such as DCA (ter Braak and Šmilauer 2015), can be used to quantify spatiotemporal variation and to identify ‘outliers’, i.e., samples that are very different to other spatial or temporal replicates.</p>
---	---	--	---

4	Determine the optimal dissimilarity measure and whether data should be standardised	<p>The dissimilarity measure selected depends on the response variable type (presence data, relative abundance data, etc.)</p> <p>This choice of appropriate distance measure and/ or standardisation determines the degree of emphasis of sample richness, composition or relative abundances in the resulting distance matrix, and thus may also relate to the research question (Anderson et al. 2011).</p>	<p>Possible dissimilarity measures to compare samples should be selected using criteria listed in sources such as Anderson et al. (2011), Legendre & Legendre (2012) and Legendre & Cáceres (2013).</p> <p>Principal coordinates analysis or non-metric multidimensional scaling (NMDS) of the different dissimilarity matrices can be compared. The distance measure that gives the greatest spread of points in a few dimensions and, therefore better discrimination among the samples, should be selected. If NMDS is used, the stress values associated with the different distance measures can be compared, because lower stress values show that the ecological distances between samples are being better represented in ordination spaces.</p>
---	---	--	--

1637 **Table 2:** Descriptions of the outputs obtained from analysis methods for temporal change in community composition with example from the
 1638 literature. LCBD refers to ‘Local Contributions to Beta Diversity’ and SCBD refers to ‘Species Contributions to Beta Diversity’. * Venn
 1639 diagrams and time cores are specifically useful for detecting taxon replacement over time, and are described in the section “Visualising temporal
 1640 patterns”.

Method (text section)	Description of outputs	Key paper(s) using each method
Beta diversity decomposition (nestedness, turnover) (D 1.2)	Beta diversity values for nestedness and turnover that can be used in subsequent graphs (e.g., time lag) or analyses of explanatory variables.	Baselga & Orme (2012)
Beta diversity decomposition (SCBD, LCBD) (D 1.2)	List of relative contributions of taxa to the overall compositional variation. Matrix of LCBD values showing the relative contribution of each sample-time combination to the overall compositional variation. Values can be used in subsequent graphs or analyses of explanatory variables.	Legendre & De Cáceres (2013); Legendre & Gauthier (2014); Lamy et al. (2015)

Beta or zeta diversity time-lag graph (D 1.2 or 2.0)	Graph showing the relationship between temporal diversity and time lag (distance in time). In the case of zeta and multi-site dissimilarity, time-lag graphs are specific to each number of samples in the subset (e.g., zeta order).	Hui & McGeoch (2014); McGeoch et al. (2017)
Beta or zeta time lag regression (D 1.2)	Regression statistics indicating the significance of the relationship between temporal diversity and time lag (distance in time). In the case of zeta and multi-site dissimilarity, time-lag analyses are specific to each number of samples in the subset (e.g., zeta order).	McGeoch et al. (2017)
Community trajectory analysis (D 1.7)	Principal coordinates analysis diagram with trajectories showing the relative change in species composition of samples through time. Trajectory statistics, including shape, size and direction are also obtained.	De Cáceres et al. (2019)

Multivariate regression modelling (D 5.0)	Multivariate regression models quantifying the relative effects of species (or species groups) interactions, environmental covariates, spatial structure, and observation error on relative abundances through time. Effects can also be output as estimated trajectory plots. Options exist to produce metrics of community stability.	Hampton et al. (2013)
Mean dissimilarity (D 1.1)	Mean dissimilarity values for data subsets. Values can be used in subsequent graphs, analyses or maps.	Collins (1992)
Mean dissimilarity as response in regression (D 1.1)	Regression statistics indicating the significance of the relationship between mean dissimilarity and explanatory variables, which includes time if dissimilarity is calculated for different spatial samples, or will only include change in non-time explanatory variables, if dissimilarity is calculated across time.	Collins (1992); Barros et al. (2014)

Moran's eigenvector maps (MEMs) with db-RDA (D 1.5)	MEMs results in a matrix of uncorrelated temporal variables that characterize scales of temporal variation in composition that can be used in the db-RDA. The RDA results in a table of ordination statistics, an RDA ordination diagram, and if paired with variance partitioning, a Venn diagram showing the amount of variance explained by each component.	Legendre & Gauthier (2014)
Ordination (constrained) (D 1.4)	Ordination diagram and table of statistics (see the text section 'Ordination' for more details and other possible outputs).	Anderson & Willis (2003); http://ecology.msu.montana.edu/labds/R/
Ordination site scores in glm or non-linear regression (D 1.4a)	Regression statistics indicating the significance of the relationship between species composition and explanatory variables.	Day & Buckley (2013)
Pivot days (D 1.6)	A list of time points are identified at which significant shifts in composition have occurred.	Lellouch et al. (2014)

Principal response curves (D 1.4b)	Principal response curve diagram comparing trajectories in compositional space of different sites or samples over time relative to the baseline or control.	Auber et al. (2017)
Raw dissimilarity values (D 1.1)	Various, depending on the method. For example, the dissimilarity-overlap method results in a graph of pairwise dissimilarity values against the proportion of individuals that taxonomically overlap; overlaid on this is a non-parametric, non-linear regression line, the slope of which indicates whether or not samples are changing in similar ways or not; this is compared to a null hypothesis of zero slope.	Bashan et al. (2016)
Synchrony (D 3.0)	A single value representing the degree to which taxa are changing in a similar way over time, calculated for a given sample by species matrix across a set of times. Can be compared for different time windows or different subsets of samples (e.g., experimental treatments).	Loreau & Mazancourt (2008); Gouhier & Guichard (2014)

Temporal coherence (D 1.4a)	A single value representing the degree of consistency in temporal change among samples for the dataset or subset considered (see the text section ‘Ordination’ for more details).	Rusak et al. (1999); Angeler & Johnson (2012)
Turnover rates (D 4.0)	Percent change in species composition between time points.	Diamond (1969)
Venn diagrams, overlap and time cores*	Venn diagram showing the number of taxa shared and unshared between two or more subsets of samples. Lists of shared and unshared taxa. Time cores results generate lists of taxa that occur consistently through time for a set of samples.	Vallès et al. (2014); O’Sullivan et al. (2015)
Zeta diversity (D 2.0)	Zeta diversity values for each zeta order (number of samples in subsets) that can be used in subsequent graphs (e.g., zeta diversity vs. order) or analyses of explanatory variables.	Hui & McGeoch (2014); McGeoch et al. (2017)

1642 **Table 3:** Data visualisation methods and considerations for their use in temporal community dynamics studies.

1643

Methods	Description and example references	Considerations
Ordination axis samples scores	Symbols (e.g., Noll et al. 2005), envelopes (e.g., Rayner et al. 2009), or trajectories (e.g., Rojas-Sandoval et al. 2012) can be added to ordination diagrams to display the relative compositional change among temporal samples.	Can be used to show convergence, divergence, or in/stability of sets of samples. Useful for many samples and taxa. Ordination sample scores can also be used in other graphs or maps to display spatiotemporal compositional variation.
Relative abundance distributions	Plot the relative abundance of all taxa in the community (Lambshead et al. 1983, Stead et al. 2003, Ruhl 2008).	Shows changes in evenness over time where separate relative abundance distributions are plotted for each time. Useful only for small numbers of comparisons, e.g., two or three sampling times.

Venn diagrams	Displays the overlap in species composition among samples (e.g., O’Sullivan et al. 2015).	Suitable for visualising taxon replacement while specifically identifying the taxa driving compositional shifts. Useful only for small numbers of temporal replicates.
Time cores	Identify ‘core’ groups of taxa that remain present over time (Vallès et al. 2014).	Useful for visualising synchronicity in compositional shifts.
Heat maps	A grid representing spatial and/or temporal samples where cells are coloured by a gradient representing relative abundances, ordination sample scores, or compositional change (e.g., Pedersen et al. 2017, Pereira et al. 2017).	Useful for representing change between time points across sample space, even in the case of complex patterns.
Geographic maps	Grid cells or other geographic polygons displaying numeric values representing relative abundances, ordination sample scores, or compositional change (e.g., Clavero and Garcia-Berthou 2006).	Useful for representing change between time points where variation in geographic space is also important. Useful for relatively few time points.

Abundance graphs	Bar graphs, boxplots, scatterplots, or pie charts showing relative abundance of different taxa at different times (e.g., Xiong et al. 2014, Chapman and McEwan 2016, Franca et al. 2016).	Can show complex temporal dynamics such as nonlinearities in species' responses and synchronicity among different taxa, but useful only for a small number of taxa.
Relative abundance tables	Relative abundance or percent dominance values across time (e.g., Shertzer et al. 2009).	Useful only for a small number of taxa and spatiotemporal samples.
Photographs of gels or electropherograms	Show the relative frequencies of characteristic taxon identifiers (e.g., Shi et al. 2016).	Useful only for a small number of taxa and spatiotemporal samples.

1644

1645

FIGURES

Figure legends

Fig. 1. Number of studies identified by our search criteria from each continent. Numbers of studies were: Global ($n = 5$), Oceanic ($n = 5$), Africa ($n = 20$), Antarctica ($n = 6$), Asia ($n = 47$), Australia ($n = 18$), Europe ($n = 166$), North America ($n = 161$), Oceania ($n = 8$) and South America ($n = 39$). Darker colours indicate more studies have been conducted within those continents, including Oceania. Pie charts show the taxonomic focus of study data from each continent as being on (●) plants, (●) vertebrate animals, (●) invertebrate animals, (●) microeukaryotes, (●) prokaryotes, (●) viruses, or (●) mixed taxonomic groups.

Fig. 2. The frequency of use of different temporal community dynamics analysis methods across all reviewed studies. Specifically: (a) the number of different analyses applied in each study; (b) the percent of uses of each analysis type across all studies and times. Descriptions of analysis types are detailed in either S1 or S2 Text.

Fig. 3: The relative distribution of the mean number of independent temporal samples across different research aims and temporal analysis methods. ‘Minor methods’ include compositional pivot days, frequency change, Markov chain modelling, multiplicative change, Moran's eigenvector maps (MEMs), nestedness analysis, and synchrony. Descriptions of analysis types are detailed in Section D ‘Road map for analysing temporal community dynamics’ in the text. Research aims vary from understanding compositional dynamics at fine, medium and coarse temporal scales, to investigating the effect on species composition over time of species interactions, environmental conditions, and ecological disturbance, to understanding successional change in communities.

Fig. 4. Conceptual figure illustrating the different patterns in (a) temporal community dynamics for a single sample unit expected under different hypotheses of community change, visualised as how pairwise dissimilarities (D), e.g., Bray-Curtis, would change over different time lags: (b) the null hypothesis of no change ('persistence'), (c) 'directional' change, and (d) 'cyclical change'/'periodicity'. Points represent individual pairwise comparisons of across eight sampling occasions and the error envelope is the standard deviation (absent for the largest lag because there is only one comparison at that scale).

Fig. 5. Conceptual figure illustrating the different patterns in (a) temporal community dynamics for a set of sample units (Sample units $A-C$) expected under different hypotheses of community change, visualised as how pairwise dissimilarities (D), e.g., Bray-Curtis, would change over eight sampling occasions (Time 0-8): (b) the null hypothesis of no change ('persistence') or 'directional', whereby all sampling units are moving at the same direction and speed in compositional space, (c) 'divergence', (d) 'instability', and (e) 'convergence' (blue line)/ 'homogenisation' (red line). Points represent individual pairwise comparisons for three sample units at eight sampling occasions, and the error envelope is the standard deviation of the three values.

Fig. 6. Conceptual figure illustrating the different patterns in (a) temporal community dynamics for a set of experimental sample units in two treatments (X and Y) expected under different hypotheses of community change, visualised as how pairwise dissimilarities calculated between treatment pairs (D), e.g., Bray-Curtis, would change over eight sampling occasions (Time 0-8): (b) a 'press perturbation' that occurs at Time 0 or a (c) a 'pulse perturbation' that creates a 'pivot point' or 'threshold' that occurs at Time 3. Both pulse and press perturbations can generate either a 'tipping point' (red line) or 'resilience' (blue line)

type response. Points represent individual pairwise comparisons for three sample units at eight sampling occasions, and the error envelope is the standard deviation of the three values.

Fig. 7. Conceptual figure illustrating the idea of ‘synchronicity’, whereby different functional or taxonomic subsets of a community are changing differently from other subsets (but the taxa within those respective groups are changing synchronously). In this case, (a) pairwise dissimilarities are calculated separately for the different subsets and (b) any possible differences in changes over time between the two subsets can be observed, e.g., one functional group might be becoming more homogeneous (blue), whereas another might be changing directionally (red). Points represent individual pairwise comparisons for two different functional groups (red and blue) from three sample units at eight sampling occasions, and the error envelope is the standard deviation of the three values.

Fig. 8: Changes in compositional turnover over time for data from the Porensky et al. (2016a, b) case study. Four transects were subjected to one of six grazing treatments and were sampled annually from 2007 to 2014. Note that “exclosure” treatment is the control (no grazing). The different lines represent different transects.

Fig. 9: Local contributions to beta diversity (LCBD) for each transect at each sampling time for data from the Porensky et al. (2016a, b) case study. Plot shows the relative contributions of the different transects and times to the overall beta diversity. Symbol size is proportional to the LCBD value and significance is indicated by the red, outlined circles; blue circles are not significant. Transects had a grazing treatment and were established in one of two replicate pastures (1 or 2) and were either north-facing (f) or south-facing (s). Grazing treatments

were: none (exclosure), light, moderate, heavy, heavy-to-light (lightly grazed since 2007; ‘new light’) and heavy-to-none (new exclosure since 2007).

Fig. 10: Principal response curve analysis of data from the Porensky et al. (2016a, b) case study. Results show that treatments differed in plant species composition from the control (exclosure; grey line) but the differences between treatments were inconsistent over time. Treatments were: new exclosure (since 2007; dark blue), continuous light (red), new light (since 2007; light blue), continuous moderate (green), and continuous heavy (black). Four particularly influential species where either lower (BOGR = *Bouteloua gracilis*) or higher (PASM = *Pascopyrum smithii*, STCO = *Hesperostipa comata*, and BRTE = *Bromus tectorum*) in abundance than in the control (black horizontal lines represent the abundance of each species relative to the control).

Fig. 11. Zeta diversity analysis, based on the number of species shared by a set of temporal samples from the Barro Colorado Island case study. Results show that the woody forest community in the 50-ha forest census plot on BCI has been significantly more stable over time than random expectation, particularly at large spatial scales where richness is higher. Maps show for 1 - zeta diversity calculated for quadrats of increasing sizes (20×20 m, 50×50 m, 125×125 m, 250×250 m and 500×500 m) and for between 2 and 7 orders of zeta (the number of temporal samples compared): (A) the observed values, (B) observed - expected values under a null model of random change in species composition in quadrats over time, and (C) significance of the observed values. Results are given as 1 - zeta so that they are comparable to the turnover results in that higher values represent higher dissimilarity in species composition.

Fig. 12. Turnover analysis results for the seven time intervals for the Barro Colorado Island forest community. Results show that the proportion of species turning over each measurement interval was significantly lower than under random expectation, particularly at large spatial scales where richness is higher and composition is more stable: Maps show proportion of species turnover calculated for quadrats of increasing sizes (20×20 m, 50×50 m, 125×125 m, 250×250 m and 500×500 m) across the seven time intervals: (A) the observed values, (B) observed - expected values under a null model of random change in species composition in quadrats over time, and (C) significance of the observed values.

Fig. 13. Principal coordinates analysis of 74 marine bacterial OTUs across 62 temporal samples from the Gilbert et al. (2012) case study. Results show no visible consistency in compositional change by season (coloured points) within or between years for the single site analysed. Black lines link samples taken consecutively within years.

Fig. 14. Time lag analysis of 74 marine bacterial OTUs across 62 temporal samples at one site off the coast of Plymouth, United Kingdom from the Gilbert et al. (2012) case study. Results show high variability in pairwise Bray-Curtis dissimilarity values with slight divergence occurring over time. The blue line is a linear regression line. The high variability suggests that there is not consistent, directional temporal change at this site.

Fig. 15. Zeta diversity analysis of the 13 marine bacterial OTUs that were most variable in abundance across 62 temporal samples from the Gilbert et al. (2012) case study. Results show that spring samples were more diverse than winter or summer samples. Zeta order represents the number of temporal samples compared and so each point on the graph shows the mean number of OTUs shared among these different numbers of temporal samples.

1770

1771 Fig. 16. Change in the mean Bray-Curtis dissimilarity values calculated for temporal samples
 1772 ($n = 62$) grouped by season from the Gilbert et al. (2012) case study. Results are for the 13
 1773 most variable marine bacterial OTUs at the one site analysed, showing cyclical community
 1774 dynamics. The blue line is a polynomial regression line with 16 terms. Multiple comparisons
 1775 from an analysis of variance showed that samples taken in spring were significantly less
 1776 similar to other spring samples than were samples taken in summer, autumn, or winter ($P <$
 1777 0.001). Samples taken in summer were significantly less similar to other summer samples
 1778 than were samples taken in autumn ($P < 0.05$).

1779

1780 Fig. 17. Potential methods that are appropriate for testing different predictions of temporal
 1781 patterns in community dynamics indicating the required temporal sample size and whether:
 1782 environmental variables can be included in the analysis (Env.); the analysis can be made
 1783 explicitly spatial to account for temporal samples being clustered in different sites by
 1784 including spatial variables (Spatial); or the role of individual taxa are identified in the
 1785 analysis (Taxa), and the type(s) of significance testing (Sig.) that are used for the analysis
 1786 (Rnd. = randomisations, Reg. = regression statistics, None = significance testing is not
 1787 normally applied, Mtl. = Mantel test). Note that the null hypothesis for each of these
 1788 predictions is that there is no change over time in either taxon presence or abundance
 1789 ('persistence'). Cell shading indicates the number of temporal samples in the time series for
 1790 which each analysis method is appropriate (note that some hypotheses require more temporal
 1791 samples than others): No shading—method applies well to time series of any length; cross-
 1792 hatched pattern—method only applies well for few measures in a time series (< 5 temporal
 1793 samples); Light grey—only applies well for ≥ 5 temporal samples; Mid-grey—only applies
 1794 well for ≥ 10 samples; Darkest grey—only applies for ≥ 20 samples, but many more than 20

1795 spatial and temporal replicates may be required, depending on the number of explanatory
1796 variables in the regression model; note that a time series of at least 20 temporal samples are
1797 required for calculating Moran's eigenvector maps (MEMs). (✓) indicates that although
1798 these statistics are not explicitly spatiotemporal analyses, they can be calculated for
1799 spatiotemporal subsets of sample units and therefore, such groupings could be used to
1800 compare variation in both space and time. + Requires a strong gradient in composition that is
1801 well represented by the ordination axis scores for samples. ++ Weak inference from species'
1802 positions on biplot. # Zeta diversity modelling can be combined with generalised
1803 dissimilarity modelling (Ferrier et al. 2007) to assess the effect of environmental variables. ##
1804 PRC requires a 'baseline' or a 'control' treatment level against which other samples are
1805 compared through time. Beta diversity decomposition refers to nestedness, turnover, and
1806 'Local Contributions to Beta Diversity' (LCBD). * Venn diagrams and time cores are
1807 specifically useful for detecting taxon replacement over time, and are described in the section
1808 "Visualising temporal patterns". Numbers in parentheses next to each method refer to the text
1809 section that gives a detailed description of each method.
1810

Figure 2

(a)

(b)

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

1851 Figure 9

1852

1853

1854 Figure 10

1855

1856

1857 Figure 11

1858 (A)

1859

1860 (B)

1861

1862 (C)

1863

1864

1865 Figure 12

1866 (A)

1867

1868 (B)

1869

1870 (C)

1871

Figure 13

1876 Figure 14

1877

1878

1879

Figure 15

1885 Figure 16

1886

1887

1888 Figure 17.
1889

Method (text section)	Also interested in?			Hypotheses of temporal community dynamics					
	Env.	Spat.	Taxa	Directional Succession Invasion Divergence Convergence Homogenisation Degradation	Periodicity Cyclical	Instability Stabilisation	Perturbation Pivot point Tipping point Threshold Resilience	Synch- ronicity	Sig. test type
Mean dissimilarity (D 1.1)		(✓)		✓	✓	✓	✓	✓	Rnd.
Beta diversity decomposition (D 1.2)		(✓)	✓	✓	✓	✓	✓	✓	Rnd.
Zeta diversity (D 2.0)	#	(✓)		✓	✓		✓	✓	Rnd.
Synchrony (D 3.0)		(✓)						✓	Rnd.
Turnover rates (D 4.0)		(✓)		✓			✓		Rnd./ Reg.
Community trajectories (D 1.7)		(✓)		✓	✓	✓	✓		Rnd.
Venn diagrams and time cores*		(✓)	✓	*					None
Time lag graph or regression (D 1.3)		(✓)		✓	✓	✓	✓	✓	Mtl/ Rnd.
Pivot days (D 1.6)		(✓)					✓		Rnd.
Multivariate regression (D 5.0)	✓	✓	✓	✓	✓	✓	✓	✓	Rnd./ Reg.
Regress mean dissimilarity (D 1.1)	✓	✓		✓	✓	✓	✓	✓	Reg.
MEMs with db-RDA (D 1.5)	✓	✓	✓	✓	✓	✓	✓	✓	Rnd.
Constrained ordination (D 1.4)	✓	✓	++	+					Rnd.
Ordination site scores (D 1.4a)	✓	✓	++	+	+	+		+	Reg.
Temporal coherence (D 1.4a)		(✓)		+				+	Rnd.
PRC (D 1.4b)##	✓	✓	✓	+	+			+	Rnd.

← Increasing sample size required ←
↓

1890

Appendix S1

Figure S1. Proportion of studies identified using our search terms that were published each year until September 2017. No studies prior to 1991 were captured by our search terms. The total number of studies in each year is the value above the bars.

Figure S2. (A) The numbers of small and large spatial scale studies by habitat category. Small spatial scale (blue bars) studies were considered to be those carried out at micro, point sample, and local scales, while large spatial scale studies (purple bars) were those at regional, continental, or global scales. (B) The proportion of studies of different taxa within categories defined by the temporal grain (black circles) and extent (purple dots). Temporal grain is the minimum time between sampling events within the study and the temporal extent is the maximum time between sampling events within the study.

(A)

(B)

Figure S3. (A) The proportion of uses of each type of analysis across all sampled publications between 1990 and 2017. ‘Minor methods’ include compositional pivot days, frequency change, Markov chain modelling, multiplicative change, Moran's eigenvector maps (MEMs), Nestedness analysis, and synchrony. (B) The popularity of the different ordination methods used across the 27 years. ‘Minor methods’ are: dbRDA, DCCA, pCCA, pRDA, Procrustes, RA and multiple co-inertia analysis. Descriptions of analysis types are detailed the S1 Text supplementary material.

Figure S4. The distribution of studies across the different research aims. Research aims vary from understanding compositional dynamics at fine, medium and coarse temporal scales, to investigating the effect on species composition over time of species interactions, environmental conditions, and ecological disturbance, to understanding successional change in communities. Scales of temporal dynamics were classified as fine (days), medium (weeks, months, or seasons), and coarse (annual, years, or decades).

Appendix S2. Other methods encountered in the literature review, not explained in the text

UNIVARIATE LINEAR METHODS

Univariate linear methods, including analysis of variance (ANOVA), analysis of covariance (ANCOVA), repeated measures ANOVA (rmANOVA), simple correlation (e.g., Pearson's and Spearman's correlation coefficients), and regression (linear regression, generalised linear models, generalised linear mixed effects models, logistic regression, *etc.*), are used to examine temporal changes in a univariate response (dependent) variable. The independent variable is either time, or an environmental variable that changes over time, and the response variable usually is the (relative) abundance of a focal taxon or functional group.

Univariate linear methods are applied most frequently to community dynamics by constructing separate models for each species. These methods of assessing compositional dynamics are limited because multiple analyses, conducted in parallel on selected taxa, are required to make inferences regarding composition. For even modest numbers of taxa, these multiple analyses can become unwieldy and require *P*-value corrections for multiple tests (Benjamini and Hochberg 1995, Zar 1996). For example, Bugalho et al. (2013) used repeated-measures mixed-effects modelling to assess the effects of deer exclusion over time on the abundance of seedlings and saplings of forest tree species. In this case, the univariate modelling approach was usefully applied because fewer than ten species were analysed. Alternatively, abundance can be analysed for all species in the same model as long as there are enough degrees of freedom in the model to include species as a fixed or random effect term (e.g., Buckley and Freckleton 2010); see Harrell (2001) for a rule of thumb stating that we should have at least ten independent samples for each predictor variable in our model. Generalised linear models are more flexible in that they can handle designs with unbalanced data and nested sampling (e.g., plots within sites) (Bolker et al. 2009, Zuur et al. 2009). R code is provided in Appendix S4.

CLUSTER ANALYSIS (A DISSIMILARITY BASED METHOD)

Cluster analysis is a catch-all term applied to dissimilarity-based methods that either group samples together (agglomerative methods) or split all samples into sub-groups (divisive methods); the clustering is based on the distances between clusters (Legendre and Legendre

2012). There are two approaches for using cluster analysis to assess community dynamics: (1) identifications of transitions between clusters over time, and (2) examination of differences of clusters of repeated observations. In the first approach, a separate cluster analysis is computed for each temporal sample of community data (e.g., Day and Buckley 2013). In the second approach, all samples taken at different times are included in a single cluster analysis (e.g., Araujo et al. 2008). If compositional shifts are small relative to the total compositional variation in the dataset, the first approach will not be able to detect temporal differences. However, if compositional shifts are relatively large, it yields a clear way of showing shifts in ‘community types’: groups of samples of similar composition that can be characterised using a method such as indicator species analysis (e.g., Day and Buckley 2013). The first method can potentially be applied to a larger number of time points if all that is desired is to report the percent of sites that changed cluster groups each time. For example, cluster analysis of 34 phytoplankton species in Laoshan Bay, China revealed distinct groups of species occurred at different times over a six-month period (Xu et al. 2010).

TEMPORAL STABILITY (A.K.A. COEFFICIENT OF VARIATION)

The coefficient of variation (CV) is used to measure temporal stability of the abundance or biomass of an individual taxon, or group of taxa, across all times (not space); this is not a multivariate method (Grossman et al. 1990, Brown and Lawson 2010, Hector et al. 2010). The CV is the ratio of the standard deviation to the mean, and often is multiplied by 100 to obtain a percentage. The mean of the CV values for all taxa (e.g., abundance or biomass) is used as an aggregate measure of temporal stability for a whole community; smaller values imply greater stability. Because the CV is a ratio, changes in community composition can be compared even if there are differences in total abundances among taxa or in taxon richness among taxon groups. Such values are often presented in tables or the text as a measure of variation or are sometimes used as a response variable against other (spatial) variables of interest (e.g., Doxa et al. 2012). For example, CV can be calculated for disturbed and undisturbed communities to assess if the disturbance has resulted in a more temporally variable community. Grossman et al. (1990) categorised particular values of CV to assess stability of freshwater fish communities: a community was ‘highly stable’ if it had a CV < 25%, ‘moderately stable’ if $25\% < CV < 50\%$, ‘moderately fluctuating’ if $50\% < CV < 75\%$ or ‘highly fluctuating’ if $CV > 75\%$. The statistical significance of CV values could be obtained by comparing the observed values to those from communities generated using

randomisations under sensible null models. Note that temporal CV is necessarily be measured on a per site basis to exclude spatial variation.

FREQUENCY CHANGE

Changes in the frequency of individual taxa within a community over time can be assessed by subtracting their frequency at each sampling occasion from a prior sampling occasion; this is not a multivariate method. Permutation tests can be used to test the significance of these changes in frequencies (Kapfer et al. 2011). Community samples (with their given taxon frequencies) are randomised among sampling times and the change in taxon frequency recalculated with each permutation. A significant change in the observed taxon frequency is inferred if the observed change is greater than the 95th percentile change from the permutations. Permutations should be spatially restricted so that only samples from homogeneous habitat types are compared; otherwise the chance of spuriously detecting change is greatly increased. The outcome of these tests is a list of taxa that had significant increases or decreases in frequencies, and the amount, between sampling times. For example, Kapfer et al. (2011) used permutation tests to assess changes in frequencies of 85 plant species from two time points 50 years apart. This method could be used on data from multiple sampling occasions, but would quickly become cumbersome with many temporal observations. It can only be used to compare individual taxa over time, does not apply to rare taxa (taxa need to occur in enough sites to make the calculations tractable), and results in a (potentially) large table of taxa with significant (or not) changes, therefore, it has limited applicability to taxon-rich datasets where changes in rare taxa are of interest.

MULTIPLICATIVE CHANGE

The change in percent cover of a taxon, or group of taxa, at a single site can be used to obtain a ‘growth rate’ using a linear regression equation; this is not a multivariate method (Debinski et al. 2010, Collins and Xia 2015). This rate of change can then be used in a regression or ANOVA to assess predictors of change for each species (‘multiplicative change’). For example, different sites or sets of sites, can be compared in their relative change in the percent cover of different functional groups (Debinski et al. 2010), as long as communities consist of taxa with similar ecology and life-histories, e.g., grassland plant communities in different moisture regimes. This method appears to be most useful for studies where (1) there is little temporal replication, or at least where differences among temporal replicates are not

of interest, because the growth rate is calculated using only two temporal samples, and (2) few species, species groups, are of interest because a different analysis must be performed for each. However, it would be possible to calculate the change in percent cover for different time periods within the overall study, if one was interested in investigating finer-scale trends.

MARKOV CHAIN MODELLING

Markov chain modelling uses a series of transition probabilities (a ‘Markov chain’), which describe how the occupancy of a taxon or functional group in a site changes from time t to time $t + 1$ (Hill et al. 2004). The user varies these probabilities in space and/or time to create a set of different hypothesis-driven models that predict the observed occupancy. The model set can then be evaluated against the observed data using log-linear modelling and AIC model selection to determine the influential and significant processes on taxon site occupancy through time (Hill et al. 2004). Community dynamics usually are examined by comparing predicted changes in individual taxa, making the method easier to apply to communities of few taxa (or groups of taxa). Because models often are parameterised using the modelled data, dependence of samples needs to be dealt with through sampling design or post-hoc statistical analyses (Hill et al. 2004). This method is most appropriate for sessile communities where patch occupancy is easy to replicate and define (e.g., Hill et al. 2004, Pueyo and Beguería 2007). It best describes directional dynamics, such as secondary succession (e.g., Baasch et al. 2010, Jiménez-Franco et al. 2011) and it is relatively computationally difficult. However, it makes detailed, taxon-level predictions that can be useful for the analysis of experimental studies where researchers want to explore specific predictions from multiple scenarios, such as taxon extinctions (e.g., Wootton 2004) or restoration manipulations (Baasch et al. 2010).

MACHINE LEARNING AND DATA MINING METHODS

Machine learning is branch of computer science that deals with the development of learning algorithms that are used to explore large, multivariate datasets, and has the primary aim of generating accurate, predictive models (Carbonell et al. 1983). Common machine-learning methods include artificial neural networks and decision-tree methods which can be used with either univariate or multivariate response variables, and Bayesian versions can incorporate prior knowledge (Cutler et al. 2007, Olden et al. 2008). Whilst the field of machine learning focusses on making predictions from data, many of its methods also are used in a data-mining

capacity to discern spatiotemporal patterns from complex and noisy datasets and can be used for classification, regression, clustering, and data reduction (Olden et al. 2008).

Machine-learning methods are attractive because they do not have many of the restrictive assumptions, such as linearity or data normality, that underpin many traditional statistical methods (Olden et al. 2008). Thus, machine-learning methods are becoming increasingly popular for exploratory and predictive analyses of community dynamics (e.g., Kampichler et al. 2010). For example, Vallès et al. (2014) used self-organising maps, an implementation of artificial neural networks (Kohonen 1982), to show that complex microbial communities sampled from infants' guts formed distinct clusters at each of five sample times (based on relative abundance data), suggesting that successional processes were occurring within the first year of these infants' lives. Sheaves et al. (2010) used a multivariate classification and regression tree analysis (De'ath 2002) to investigate spatial and temporal differences in fish assemblages; results showed that despite considerable spatial variation in community structure among different estuaries, there nonetheless were clear patterns of temporal change in species' abundances and assemblage composition. These methods are very flexible and can handle both univariate and multivariate input datasets with large numbers of predictor variables, including space and time, allowing simultaneous investigation of spatial and temporal community dynamics. Standardised approaches for characterising different types of temporal patterns and causal processes in community datasets will hopefully emerge from further research. R code is not provided, but R packages are available (the MultivariateRandomForest package: Rahman 2017, and rpart package: Therneau and Atkinson 2018).

NETWORK ANALYSIS

Ecological network analysis enables the investigation of the complex associations among species, and among species and their environment, through space and time (Newman 2010, Blonder et al. 2012). This method recognises that an ecological community is a complex biological system comprised of interconnected units whose associations can be modelled mathematically using constructs such as vertices (representing taxa) and edges, which are the connections between the vertices (representing ecological interactions) (Blonder et al. 2012). In the context of investigating community dynamics, patterns of species' associations and community memberships are most typically visualised as a topological network diagram created from taxon abundance data collected at multiple time points (*sensu* "time-aggregated

networks” (Blonder et al. 2012). From the network topology, a variety of statistics can be calculated that measure variability in community composition and interactions at particular times, which can then be compared. This method is most applicable for multi-trophic communities or where interaction networks can be clearly defined for the community; it is less useful where a single trophic level is of interest and interactions are not quantified, e.g., a plant community, but see Cram et al. (2015) for an example of spatiotemporal microbial networks. In R, the package ‘timeordered’ is the easiest way to implement a network analysis of community dynamics data. R code is not provided.

REFERENCES CITED

- Araujo, H. M. P., D. A. Nascimento-Vieira, S. Neumann-Leitão, R. Schwamborn, A. P. O. Lucas, and J. P. H. Alves. 2008. Zooplankton community dynamics in relation to the seasonal cycle and nutrient inputs in an urban tropical estuary in Brazil. *Brazilian Journal of Biology* 68:751–762.
- Baasch, A., S. Tischew, and H. Bruelheide. 2010. Twelve years of succession on sandy substrates in a post-mining landscape: a Markov chain analysis. *Ecological Applications* 20:1136–1147.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57:289–300.
- Blonder, B., T. W. Wey, A. Dornhaus, R. James, and A. Sih. 2012. Temporal dynamics and network analysis. *Methods in Ecology and Evolution* 3:958–972.
- Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* 24:127–135.
- Brown, B. L., and R. L. Lawson. 2010. Habitat heterogeneity and activity of an omnivorous ecosystem engineer control stream community dynamics. *Ecology* 91:1799–1810.

183 Buckley, H. L., and R. P. Freckleton. 2010. Understanding the role of species dynamics in
184 abundance-occupancy relationships. *Journal of Ecology* 98:645–658.

185 Bugalho, M. N., I. Ibáñez, and J. S. Clark. 2013. The effects of deer herbivory and forest type
186 on tree recruitment vary with plant growth stage. *Forest Ecology and Management*
187 308:90–100.

188 Carbonell, J. G., R. S. Michalski, and T. M. Mitchell. 1983. An overview of machine
189 learning. Pages 3–23 *Machine Learning: An artificial intelligence approach*. Tioga
190 Publishing Company, Palo Alto.

191 Collins, S. L., and Y. Xia. 2015. Long-term dynamics and hotspots of change in a desert
192 grassland plant community. *The American Naturalist* 185:E30–E43.

193 Cram, J. A., L. C. Xia, D. M. Needham, R. Sachdeva, F. Sun, and J. A. Fuhrman. 2015.
194 Cross-depth analysis of marine bacterial networks suggests downward propagation of
195 temporal changes. *The ISME journal* 9:2573.

196 Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler.
197 2007. Random forests for classification in ecology. *Ecology* 88:2783–2792.

198 Day, N. J., and H. L. Buckley. 2013. Twenty-five years of plant community dynamics and
199 invasion in New Zealand tussock grasslands. *Austral Ecology* 38:688–699.

200 De’ath, G. 2002. Multivariate regression trees: A new technique for modeling species-
201 environment relationships. *Ecology* 83:1105–1117.

202 Debinski, D. M., H. Wickham, K. Kindscher, J. C. Caruthers, and M. Germino. 2010.
203 Montane meadow change during drought varies with background hydrologic regime
204 and plant functional group. *Ecology* 91:1672–1681.

205 Doxa, A., M. L. Paracchini, P. Pointereau, V. Devictor, and F. Jiguet. 2012. Preventing biotic
206 homogenization of farmland bird communities: The role of high nature value
207 farmland. *Agriculture, Ecosystems & Environment* 148:83–88.

208 Grossman, G. D., J. F. Dowd, and M. Crawford. 1990. Assemblage stability in stream fishes:
209 A review. *Environmental Management* 14:661–671.

210 Harrell, F. E. 2001. *Regression Modeling Strategies: With Applications to Linear Models,*
211 *Logistic Regression, and Survival Analysis.* Springer, New York.

212 Hector, A., Y. Hautier, P. Saner, L. Wacker, R. Bagchi, J. Joshi, M. Scherer-Lorenzen, E. M.
213 Spehn, E. Bazeley-White, M. Weilenmann, M. C. Caldeira, P. G. Dimitrakopoulos, J.
214 A. Finn, K. Huss-Danell, A. Jumpponen, C. P. H. Mulder, C. Palmborg, J. S. Pereira,
215 A. S. . Siamantziouras, A. C. Terry, A. Y. Troumbis, B. Schmid, and M. Loreau.
216 2010. General stabilizing effects of plant diversity on grassland productivity through
217 population asynchrony and overyielding. *Ecology* 91:2213–2220.

218 Hill, M. F., J. D. Witman, and H. Caswell. 2004. Markov chain analysis of succession in a
219 rocky subtidal community. *The American Naturalist* 164:E46–E61.

220 Jiménez-Franco, M. V., J. E. Martínez, and J. F. Calvo. 2011. Territorial occupancy
221 dynamics in a forest raptor community. *Oecologia* 166:507–516.

222 Kampichler, C., R. Wieland, S. Calmé, H. Weissenberger, and S. Arriaga-Weiss. 2010.
223 Classification in conservation biology: A comparison of five machine-learning
224 methods. *Ecological Informatics* 5:441–450.

225 Kapfer, J., J.-A. Grytnes, U. Gunnarsson, and H. J. B. Birks. 2011. Fine-scale changes in
226 vegetation composition in a boreal mire over 50 years: Fine-scale changes in boreal
227 mire. *Journal of Ecology* 99:1179–1189.

228 Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. *Biological*
229 *Cybernetics* 43:59–69.

230 Legendre, P., and L. F. J. Legendre. 2012. *Numerical Ecology.* Third Edition. Elsevier,
231 Radarweg, The Netherlands.

232 Newman, M. E. J. 2010. *Networks: An introduction.* Oxford University Press, New York.

233 Olden, J. D., J. J. Lawler, and N. L. Poff. 2008. Machine learning methods without tears: A
 234 primer for ecologists. *The Quarterly Review of Biology* 83:171–193.

235 Pueyo, Y., and S. Beguería. 2007. Modelling the rate of secondary succession after farmland
 236 abandonment in a Mediterranean mountain area. *Landscape and Urban Planning*
 237 83:245–254.

238 Rahman, R. 2017. *MultivariateRandomForest: Models Multivariate Cases Using Random*
 239 *Forests*. R package version 1.1.5, [https://CRAN.R-](https://CRAN.R-project.org/package=MultivariateRandomForest)
 240 [project.org/package=MultivariateRandomForest](https://CRAN.R-project.org/package=MultivariateRandomForest).

241 Sheaves, M., R. Johnston, and R. Connolly. 2010. Temporal dynamics of fish assemblages of
 242 natural and artificial tropical estuaries. *Marine Ecology Progress Series* 410:143–157.

243 Therneau, T., and B. Atkinson. 2018. *rpart: Recursive Partitioning and Regression Trees*. R
 244 package version 4.1-12. <https://CRAN.R-project.org/package=rpart>.

245 Vallès, Y., A. Artacho, A. Pascual-García, M. L. Ferrús, M. J. Gosálbes, J. J. Abellán, and M.
 246 P. Francino. 2014. Microbial succession in the gut: Directional trends of taxonomic
 247 and functional change in a birth cohort of Spanish infants. *PLoS Genetics*
 248 10:e1004406.

249 Wootton, T. J. 2004. Markov chain models predict the consequences of experimental
 250 extinctions. *Ecology Letters* 7:653–660.

251 Xu, H., G.-S. Min, J.-K. Choi, K. A. S. Al-Rasheid, X. Lin, and M. Zhu. 2010. Temporal
 252 dynamics of phytoplankton communities in a semi-enclosed mariculture pond and
 253 their responses to environmental factors. *Chinese Journal of Oceanology and*
 254 *Limnology* 28:295–303.

255 Zar, J. H. 1996. *Biostatistical Analysis*. Third. Prentice Hall, New Jersey.

256 Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. *Mixed Effects*
 257 *Models and Extensions in Ecology with R*. Springer, New York.

258

259