Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition

Franck Jabot[1*], Fabien Laroche[2], François Massol[3,4], Florent Arthaud[5], Julie Crabot[6], Maxime Dubart[7], Simon Blanchet[8], François Munoz[9], Patrice David[7], Thibault Datry[6]

[1]Université Clermont Auvergne, INRAE, UR LISC, Centre de Clermont-Ferrand, 9 avenue Blaise Pascal CS 20085, F-63178 Aubière, France

[2]INRAE, UR EFNO, Centre de Nogent-sur-Vernisson, Nogent-sur-Vernisson, France

[3]Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, SPICI group, F-59000 Lille, France

[4]Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIIL - Center for Infection and Immunity of Lille, F-59000 Lille, France

[5]Univ. Savoie Mont Blanc, INRAE, CARRTEL, 74200 Thonon-les-Bains, France

[6]INRAE, UR Riverly, Centre de Lyon-Villeurbanne, 5 rue de la Doua, 69625 Villeurbanne Cedex, France

[7]CEFE UMR 5175, CNRS - Université de Montpellier - Université Paul-Valéry Montpellier – IRD - EPHE, 1919 route de Mende, 34293 Montpellier cedex 5, France

[8]CNRS, Université Toulouse III Paul Sabatier, Station d'Écologie Théorique et Expérimentale, UMR 5321, 2 route du CNRS, 09200 Moulis, France

[9]University Grenoble-Alpes, LECA, Grenoble Cedex 9, France

[*]Corresponding author: Franck Jabot (franck.jabot@inrae.fr)

Authorship statement: all authors conceived the original idea and gave final approval for publication. FJ, FA, JC and MD performed the analyses. FJ lead the writing of the draft with contributions from all authors.

Data accessibility statement: all data are available online – see l. 673-678.

Number of words in abstract: 149.

Number of words in main text: 4999.

Number of cited references: 77.

Number of tables: 1.

Number of figures: 4.

Number of text boxes: 1.

Abstract

Although metacommunity ecology has been a major field of research in the last decades, with both conceptual and empirical outputs, the analysis of the temporal dynamics of metacommunities has only emerged recently and consists mostly of repeated static analyses. Here, we propose a novel analytical framework to assess metacommunity processes using path analyses of spatial and temporal diversity turnovers. We detail the principles and practical aspects of this framework and apply it to simulated datasets to illustrate its ability to decipher the respective contributions of entangled drivers of metacommunity dynamics. We then apply it to four empirical datasets. Empirical results support the view that metacommunity dynamics may be generally shaped by multiple ecological processes acting in concert, with environmental filtering being variable across both space and time. These results reinforce our call to go beyond static analyses of metacommunities that are blind to the temporal part of environmental variability.

Key words: beta-diversity; demographic stochasticity; dispersal limitation; environmental filtering; path analysis.

Introduction

A main goal of community ecology is to understand the determinants of species diversity at different spatial scales. Metacommunity theory has emerged as a framework to investigate the spatial distribution of species and the dynamics of spatially structured ecosystems (Leibold et al. 2004, Massol et al. 2011, Guichard 2017). Metacommunity theory has been originally proposed to include four main (not mutually exclusive) paradigms explaining the coexistence of species (Leibold et al. 2004, Shoemaker & Melbourne 2016, Fournier et al. 2017; but see also criticism of Brown et al. 2017), and synthesizing how basic processes can drive metacommunity assembly. The patch-dynamic paradigm focuses on the processes of competition, colonization and extinction in networks of patches (Hastings 1980, Tilman 1994, Calcagno et al. 2006). The species-sorting paradigm focuses on the differential responses of species to environmental heterogeneity across the landscape to explain large-scale and local coexistence as the result of environmental filters and local adaptation (Chase & Leibold 2003). The mass-effect paradigm focuses on source-sink dynamics among communities, with species potentially co-occurring in patches where they are maladapted due to the influx of dispersing individuals (Amarasekare & Nisbet 2001, Mouquet & Loreau 2003). The neutral paradigm solely considers the effects of demographic stochasticity and dispersal limitation on community dynamics, thus explaining local species co-occurrence as a stochastic process driven by species frequencies at a larger scale and immigration (Hubbell 2001). These four simplistic views of metacommunities were defined to encompass the main models and assumptions on coexistence mechanisms, both in theory and in empirical studies (Cottenie 2005, Shoemaker & Melbourne 2016, Ulrich et al. 2017).

Metacommunity paradigms and related models have mostly been used to analyse spatial patterns of metacommunity composition at a single date, assuming that metacommunities are then at a dynamical equilibrium (Logue et al. 2011, Heino et al. 2015), and often with statistical models not founded on dynamical models (but see Azaele et al. 2006). Specifically, when spatial environmental variation is hypothesized to play a role, the most common approach is to perform variance partitioning (Borcard et al. 1992, Cottenie 2005, but see e.g., Leibold and Mikkelson 2002, Ulrich et al. 2017). It consists in partitioning the observed spatial variation of community composition into spatial and environmental components (Borcard et al. 1992, Cottenie 2005, Peres-Neto et al. 2006). The effect of the spatial component is then expected to reflect a combined effect of dispersal and ecological drift, while the effect of the environmental component should summarize differential species responses to environmental variation. Such analyses of static spatial patterns of metacommunities provided insights on the processes structuring metacommunities across biomes, taxa and along environmental gradients (Cottenie 2005, Henriques-Silva et al. 2013, Heino et al. 2015). However, results on simulated datasets suggest that partitioning alone does not allow unambiguous inference of metacommunity dynamics (Gilbert & Bennett 2010, Peres-Neto & Legendre 2010).

Ecosystems and their constituent communities are highly dynamic in time (e.g., Brokaw 1985, Tscharntke et al. 2005, Malard et al. 2006, Bertrand et al. 2016), and temporal variation in community composition can impair the analysis of metacommunity diversity at a single date (Box 1). Temporal data should thus provide key information on community processes and assembly dynamics (Anderson and Cribble 1998, Magurran and Henderson 2010, Wolkowich et al. 2014, Buckley et al. 2018). To date, few studies have examined the temporal dynamics of metacommunities (Datry et al. 2016). They have mainly focused on describing spatiotemporal patterns (e.g. Soininen 2010, White et

al. 2010, Legendre & Gauthier 2014). We here argue that such limited emphasis reflects (i) a lack of a general quantitative framework to analyse temporal changes (but see e.g., Nuvoloni et al. 2016) and (ii) the scarcity of proper empirical datasets.

Nuvoloni et al. (2016) proposed to analyse the temporal turnover of community composition and to relate local turnover to environmental variables (see also Legendre 2019). We here propose to generalize this approach to metacommunities, that is to jointly analyse spatial and temporal turnovers of community composition. Indeed, using temporal signatures in metacommunity analyses is likely to reveal previously undetected environmental effects (Box 1). We therefore propose to perform path analyses that incorporate the influence of environmental, dispersal and community context on spatio-temporal turnovers of community composition. This approach presents the advantage of being sufficiently simple to elucidate the complex direct and indirect relationships among the drivers.

We propose a heuristic path model (Fig. 2). In this model, we predict that dispersal limitation and environmental filtering should entail positive correlations between community dissimilarity and, respectively, geographical distance and environmental distance (Borcard et al. 1992). Second, demographic stochasticity should entail negative correlations between mean community size and community dissimilarity across space and time, and positive correlation between temporal distance and community dissimilarity (Lande et al. 2003). Third, differences in community size should be positively linked to differences in species richness due to a general nestedness pattern of occurrence of abundant versus rare species (Srivastava & Lawton 1998), which in turn should entail greater community dissimilarity (through nestedness, see Baselga 2010). Finally, we consider that environmental distance can be correlated with temporal and geographical distances. Our heuristic understanding of spatio-temporal community dissimilarity patterns makes use of both direct and indirect relationships between explanatory variables that are themselves likely to be correlated to some extent. Path analyses therefore constitute a natural way to investigate such direct and indirect putative drivers of metacommunity dynamics (Kingsolver & Schemske 1991). Our heuristic path model is based on relationships between drivers and community composition that have been recurrently evidenced in the literature. Still, one may think of particular systems that may deviate from these general relationships and would require building of alternative heuristic path models.

We use this novel framework of spatio-temporal metacommunity analysis to analyse simulated data. We demonstrate that it enables us to detect the signature of the simulated metacommunity processes on the community patterns of beta-diversity through space and time. We then apply this framework to four real case studies. We find that spatial and temporal distances both influence dissimilarities in community composition. This effect is either direct or indirect through the spatio-temporal variations of the environmental drivers of community filtering.

Materials and methods

Assessing the path analysis framework with simulations

We devised an individual-based simulation algorithm of metacommunity dynamics in discrete time where communities are distributed across a two-dimensional grid. Our modelling choices were guided by the objectives of incorporating demographic stochasticity (using stochastic birth-death processes, Hubbell 2001), dispersal, environmental filtering (using the match between species traits

and environmental values, Gravel et al. 2006), and environmental variability in space and time. This metacommunity model is particularly suited to represent sessile or territorial organisms, for which dispersal mostly occurs at the seed, larval or juvenile stages.

Metacommunity simulator

- Regional pool

We considered a fixed regional species pool of *S* species (S=100), each species *i* having a fixed regional frequency χ_i and a fixed trait value τ_i, which corresponds to its environmental optimum. All species have the same regional frequency (χ_i =0.01) and trait values are regularly spaced between 0 and 1 (τ_i =i/100), so as to be parsimonious.

- Landscape

We considered a gridded landscape of 400 cells (20 x 20) with fixed null boundary conditions. Abiotic environmental conditions within each cell at position (i,j) in the grid are assumed homogeneous and measured with the environmental variable $E_{ij}(t)$, which can vary in time (t). This variable influences processes of adult mortality and propagule establishment in each cell. There are $J_{ij}(t)$ individuals per cell, this number varying across space and time, depending on the balance between recruitment/immigration and mortality in each cell.

- Environmental dynamics

The environmental variable $E_{ij}(t)$ in cell (i,j) at time t is decomposed into three components:

$$E_{ij}(t) = g_{ij} + a(t) + a_{ij}(t) \qquad \text{(Eq. 1)}$$

where g_{ij} represents an average environmental context in cell (i,j), a_t represents a temporal trend common to all cells, and $a_{ij}(t)$ represents a cell-specific temporal anomaly.

We consider a linear environmental gradient from the left to the right of the grid, so that g_{ij} regularly varies from $0.5 – e_1/2$ to $0.5 + e_1/2$ according to the column j of the cell (i,j), g_{ij} being constant on each column; a(t) is a triangle wave between $- e_2/2$ and $e_2/2$ with a period τ; $a_{ij}(t)$ is uniformly drawn between $- e_3/2$ and $e_3/2$ at each time step t and for each cell (i,j).

Environmental dynamics are parameterized with three parameters e_1, e_2 and e_3 to control the magnitude of the spatial environmental gradient (e1), the temporal gradient (e2), and the residual environmental variability (e3). The triangle wave function is used so as to reach a dynamical equilibrium during the burn-in period, but a large wave period is used, so that a monotonic temporal gradient is actually simulated during the recorded dynamics.

- Community dynamics

In each cell and during each time step, community dynamics is governed by four processes taking place sequentially: 1) reproduction, 2) propagule dispersal, 3) adult mortality and 4) propagule establishment. All cells are simultaneously updated.

1) Reproduction

Each individual of the community produces propagules at a constant rate r so that the number of propagules produced by each individual during one time step is a Poisson draw with parameter r (with $r \leq 1$).

2) Dispersal

Two local dispersal models are alternatively used. In these two models, a proportion (1-m) of the propagules stays in their home cell, while a proportion m disperses outside the cell. In the global dispersal model, propagules homogeneously disperse among the 400 cells of the landscape, while in the neighbour dispersal model, they only disperse to the eight neighbouring cells (uniform random draws). This neighbour dispersal process enables us to model the dispersal influence on the spatial auto-correlation of species abundances. On top of this local dispersal, additional propagules arrive from the regional pool at a constant rate I in each cell, so that the number of long-distance dispersal propagules is computed as a Poisson draw with parameter I. This spatially-implicit long-distance dispersal process enables us to maintain species diversity at the grid scale (Hubbell 2001). This simulation setting leads to realistic uneven species-abundance distributions at the grid scale, despite the even species abundance distribution in the regional pool.

3) Mortality

Each individual of species s has a local fitness $f_s(i,j,t)$ in cell (i,j) at time t, depending on the match between its trait value τ_s and the environmental variable $E_{ij}(t)$ in cell (i,j) at time t:

$$f_s(i,j,t) = 1 + A \times \exp\left[- (\tau_s - E_{ij}(t))^2/(2\sigma^2)\right] \qquad \text{(Eq. 2)}$$

where parameter A controls the strength of environmental filtering (complete maladaptation leads to a local fitness of 1 while perfect adaptation to a local fitness of 1 + A) and parameter σ controls its specificity (a relatively good local adaptation is obtained when $|\tau_s - E_{ij}(t)|$ is less than σ). This Gaussian-based equation is a rather standard way of modelling environmental filtering in community ecology (Gravel et al. 2006, Gilbert & Bennett 2010, Jabot 2010, Münkemüller & Gallien 2015, Sokol et al. 2017, Munoz et al. 2018)

The survival of adult individuals of species s is modelled at each time step t in cell (i,j) as a Bernoulli draw with probability $(1-r) \times f_s(i,j,t) / (1+A)$. It implies that the individual death probability is at least equal to r and increases as individual fitness decreases.

4) Establishment

Each cell has a carrying capacity of J individuals. The number of recruited individuals $N_r(i,j,t)$ in a cell (i,j) at time t follows a Poisson distribution with mean equal to $J-N_{ij}(t)$, where $N_{ij}(t)$ is the number of surviving adults in the cell after the mortality step. If $N_{ij}(t)$ is larger than J, then no individual is recruited. This modelling choice enables the number of individuals per cell to vary. The number of $N_r(i,j,t)$ recruited individuals follows a multinomial draw with probabilities proportional to species numbers of propagules reaching the focal cell times their local fitness $f_s(i,j,t)$.

The metacommunity is initialized with a multinomial draw of J=100 individuals from the regional pool in each cell. A burn-in period of 1,000 steps is used. Metacommunity dynamics then run for 20 time

steps. 50 cells are randomly selected at the beginning of the simulation and their composition is recorded at each time step for subsequent analyses. The code of this simulator is provided (Supplementary material S1).

Simulated scenarios

We devised 6 scenarios representing various combinations of ecological processes (Table 1). The first two scenarios implemented neutral dynamics, with either a homogeneous dispersal in the landscape or a dispersal to neighbouring cells to cause dispersal limitation. The third and fourth scenarios implemented environmental filtering along either a spatial or a temporal environmental gradient with homogeneous dispersal. The fifth and sixth scenarios implemented environmental filtering along simultaneous spatial and temporal gradients with either homogeneous dispersal or dispersal to neighbouring cells. Detailed parameter settings of the different scenarios are given in Table S2.

Path analyses

We computed Sorensen dissimilarity indices for all pairs of sampled communities. Dissimilarity of communities at a given date represents spatial dissimilarity, dissimilarity of communities at the same site but at different dates represents temporal dissimilarity, and dissimilarity of the remaining pairs represents spatio-temporal dissimilarity. We computed spatial distances (Δx), temporal distances (Δt) and environmental distances (ΔE) for each pair of communities, as well as their mean community size ($\langle J \rangle$), their absolute difference in community size (ΔJ) and in species richness (ΔS). We ran a path analysis based on the heuristic causal model (Fig. 2) with the function "sem" of the R package "lavaan" (Rosseel 2012). Since path analyses were based on distance matrices, we used the permutation-based approach developed by Fourtune et al. (2018) taking into account non-independence and allowing to test for the significance of each path. We followed a Benjamini-Hochberg procedure to adjust the significance criterion (of 1%) for multiple testing. We assessed model fit with the Standardized Root Mean Square Residual (SRMR).

Empirical datasets

We applied our conceptual framework to four empirical case studies.

- Freshwater fishes

We analysed yearly samples of freshwater fish communities from the French Office for Biodiversity database including more than 1500 sites in France (Poulet et al. 2011). We analysed the dynamics of a subset metacommunity in the Garonne-Dordogne river drainage (Fourtune et al. 2016), including 32 sites monitored each year between 1995 and 2011, and for which precise environmental variables were available. This dataset included 51 fish species, for a total of 257,393 sampled individuals. Six environmental variables were recorded for each site: elevation, slope, average temperature in January 2011, average temperature in July 2011, width of the minor riverbed, and width of the water slide. The first five variables were temporally constant, while the last variable varied from year to year. Geographical distance between sites was computed along the river using the Carthage dataset of the French National Geographical Institute. We used log-transformed distances in analyses below, but results proved qualitatively similar when using raw distances.

- Aquatic invertebrates

We compiled aquatic invertebrate communities across the Rhône river drainage in France. Benthic invertebrates were sampled on 6 sites of 11 different watersheds for a total of 66 sites. They were sampled for six months consecutively from the end of autumn to the beginning of summer for two years, 2014 and 2015, for a total of 12 sampling dates. The rivers considered are intermittent; when some sites were dry, they were not sampled at this date. Invertebrates were identified to the genus level but information was kept at the family level when no taxa were identified at the genus level for this family, resulting in a total of 231 taxa. Five environmental variables were measured at each site and sampling date: temperature, pH, conductivity, concentration in dioxygen and number of days since the last rewetting event of the watershed. We also computed log-transformed Euclidean geographic distances between sites.

- Freshwater snails

The third dataset concerns the malacological fauna – 27 species - of a freshwater ponds network in the Guadeloupe Island. 250 sites have been yearly sampled since 2001 (17 years), with species density information. Species densities were multiplied by pond area to obtain estimated species abundances in each pond, and were then log-transformed. Each site is characterized by six temporally constant environmental variables (size, depth, vegetation cover, water quality, litter and a synthetic index of hydrological and vegetation stability, see Lamy et al. 2013 for additional details), and one temporally varying but spatially constant variable (annual rainfall). We calculated log-transformed, Euclidean geographical distances among sites. Missing data and empty sites were removed prior to analyses leading to a total of *ca.* 2800 samples.

- Aquatic plants

We compiled aquatic plant communities in shallow lakes used for fish farming. These lakes are generally dried out during one year every three years. Twenty-four lakes were sampled from 2 to 7 years between 2008 and 2015, for a total of 81 sampling events and 84 sampled species (Arthaud et al. 2013). Average species cover was multiplied by lake areas to obtain estimated species abundances. Two environmental variables were used: chlorophyll a concentration reflecting water turbidity and light transmission, and the number of years since the last drying event.

Results

Analysis of simulated data

Our application of a causal modelling framework to simulated data validated our heuristic predictions and showed that the modelling framework allows reliable inferences of the ecological processes driving spatiotemporal variation in community composition for contrasted simulation scenarios. The effect of demographic stochasticity was detected in the four first scenarios with the moderate negative correlation between average community size and community dissimilarity (Fig. 3a-d). It was not detected in the last two scenarios, highlighting its minor importance compared to filtering processes in these scenarios. Demographic stochasticity was also detected with the moderate positive correlation between temporal distance and community dissimilarity in the two neutral scenarios (Fig. 3a-b). The effect of environmental filtering was detected in the four last scenarios with the positive correlation between environmental distance and community dissimilarity (Fig. 3c-f). The path analysis also correctly detected the positive correlation between spatial or

temporal distances and environmental distance in the scenarios with a spatial or temporal environmental gradient respectively (Fig. 3c-f). In the scenarios with a temporal environmental gradient, a positive correlation between temporal distance and community dissimilarity was also identified (Fig. 3d-f). In these cases, we interpret this path as a signature of the inertia of these communities to the directional temporal environmental change simulated, rather than to a signature of demographic stochasticity. Indeed, the correlations were stronger than in the neutral scenarios in these cases. This result points that the interpretation of this specific path between temporal distance and community dissimilarity should be interpreted with caution, either as a signature of demographic stochasticity if temporal distance does not impact environmental distance, or as a signature of community inertia if temporal distance does impact environmental distance. In the two scenarios with dispersal limitation, a positive correlation between geographical distance and community dissimilarity was identified (Fig. 3b,f). A sampling effect was also detected in some scenarios with the positive correlations between differences in community size and differences in community richness and in turn between differences in community richness and community dissimilarity. These sampling effects were moderate when present and did not blur the other ecologically more informative paths. Finally, one can note the weak negative correlation between geographical distance and community dissimilarity in the third scenario with a spatial environmental gradient (Fig. 3c). We consider this correlation as an artefact due to the linear modelling framework used. It is outweighed by the strong positive indirect effect of geographical distance on community dissimilarity through environmental distance (Fig. 3c).

Analysis of empirical datasets

Our statistical framework revealed very consistent patterns across case studies (Fig. 4). The influence of demographic stochasticity was evidenced in all cases (see the dashed lines from <J> to βsor). Geographic distances Δx affected community dissimilarity (βsor) in all case studies, both directly (putatively through dispersal limitation) and indirectly through environmental distances ΔE. Environmental distances ΔE influenced community dissimilarity (βsor) in all case studies. Temporal distances Δt cascaded onto environmental distances in three of the four case studies. It also directly affected community dissimilarity in half of the case studies. Finally, differences in local species richness ΔS affected community dissimilarity in all case studies, and differences in local community sizes ΔJ influenced ΔS in three of the four case studies. The later result underlined that ΔS should be taken into account when assessing the drivers of community dissimilarity.

Although we found support for the three main types of ecological drivers (demographic stochasticity, environmental variation and dispersal limitation), environmental variation was generally the strongest driver of community dissimilarity. This environmental variation was both spatially and temporally structured in three out of the four case studies (see the arrows from Δx and Δt towards ΔE). This result supports our call for an integrated spatiotemporal appraisal of metacommunity patterns.

Discussion

322 *An operational approach to analyse spatiotemporal community turnover*

We here proposed a simple approach to analyse spatiotemporal community turnover (Fig. 2). We
324 found that approach allows detecting how dispersal, demographic stochasticity and environmental filtering influence metacommunity dynamics (Fig. 3). The proposed framework is robust and general
326 since we examined strongly contrasted scenarios that all lead to path analysis results that were consistent with simulation choices. Applied to four real ecological case studies, we evidenced that
328 the environmental drivers of community composition are always spatially auto-correlated and are temporally auto-correlated in three out of the four case studies. This further supports our call for a
330 joint analysis of community turnover in both space and time.

Detecting the contributions of entangled ecological processes

332 Applied to the fish data, this spatiotemporal framework revealed that the turnover in fish community composition at yearly and regional scales is mainly driven by environmental filtering (Fig. 4a) and that
334 the environmental drivers are spatially auto-correlated ($\Delta x \rightarrow \Delta E$). We also evidenced the influence of demographic stochasticity (ecological drift) and dispersal limitation on the spatiotemporal turnover
336 component (Fig. 4a). Another main driver of community turnover is the heterogeneity in richness among local communities (ΔS), which we interpret as a nuisance variable here, since we do not have
338 specific hypotheses on what may drive this heterogeneity beyond differences in community size (ΔJ). Alternative – yet non-exclusive – explanations for the observed variability in local species richness
340 include the presence of a natural upstream-downstream gradient in species richness with more species near the outlet of the river networks (Muneepeerakul et al. 2008, Blanchet et al. 2014) and
342 the introduction of non-native species that may not be homogeneous across the river network. Our analysis reveals that such potential drivers may have a dominant effect on the overall fish
344 metacommunity structure at the regional scale.

Applied to the invertebrate data, the main driver of community turnover was also the heterogeneity
346 in richness among local communities (Fig. 4b). This may result from the fact that this dataset comprises perennial and intermittent sites, and the latter ones generally harbour species-poor,
348 original communities with taxa especially adapted to recover from disturbances (Datry et al. 2014). The other main drivers were demographic stochasticity and dispersal limitation, which may be
350 explained by the intensity of local disturbances and regional fragmentation induced by drying events. Temporal and spatial distances also have a strong effect on environmental distances, as expected for
352 intermittent rivers, as the stochasticity of drying events leads to a high spatiotemporal variability of the environment in a spatially and temporally auto-correlated way.

354 For the snail dataset, environmental variation was the main driver of community dissimilarity. Structuring environmental variables were vegetation and depth, reflecting a gradient from large,
356 deep sites, with aquatic vegetation only near the shore, to small, shallow sites often entirely covered with vegetation. Shallow sites favour some pulmonate species that live near the surface, while most
358 operculate snails are bottom-dwellers less dependent on macrophytes. Most environmental variables were spatially structured. Besides, communities were temporally structured. This may
360 reflect a mix of long-term changes (invasion of new species and decline of some resident species during the study period) and short-term stochastic extinction-colonisation dynamics that make snail

362 communities drift away from their initial states over time, within each site. For this last process, the
strong negative impact of community size on species turnover indicates that small communities
364 undergo faster temporal changes. These results point both to a strong impact of environmental
filtering and fast temporal turnover, especially in small communities, which is in line with our
366 experience with the ecology of freshwater snails.

For the aquatic plants dataset, our results point that environmental variation is a major driver of
368 community dissimilarity, in accordance with the literature (Arthaud et al. 2013). Community size is
also a significant driver of community composition. Indeed, in shallow lakes, communities are mostly
370 made of plant species that are generally abundant (competitive or ruderal species, Arthaud et al.
2012), while a few species are characterized by lower population sizes (stress tolerant species). Our
372 method evidenced that environmental filtering is both varying across space and time (Table S13). The
light stress due to phytoplankton (chlorophyll a concentration) is temporally structured, because
374 primary production varies strongly according to temperature and sunshine of the year. In contrast,
the number of years since the last drying event is spatially structured, because of the specificities of
376 the management performed at the catchment scale.

More generally, we found very consistent results across the four case studies despite the diversity of
378 taxonomic groups (plants, aquatic invertebrates, molluscs and fishes) and habitats (lakes, ponds,
perennial streams and intermittent rivers). It suggests the generality of the significance of
380 spatiotemporal variation of environmental conditions for metacommunity dynamics. Ecologists
should embrace a more dynamical view of metacommunity assembly and look beyond the
382 predominant perspective, which considers communities as assembled through temporally fixed
environmental filters. This present contribution offers a pragmatic way forward in this direction.

384 *Applying the proposed framework to metacommunity data*

The proposed framework requires temporal data of metacommunity composition and temporal
386 environmental variables that are thought to be influential for the system studied. If these
environmental data are only available at the metacommunity scale, a different path model may still
388 be used without the arrow between Δx and ΔE. Since the approach is exploratory, it does not require
a minimal amount of sampled dates nor of sampled locations (beyond 2) to be operational. In the
390 studied datasets, the number of sampled dates varied from 2 to 17, while the number of sampled
locations varied from 24 to 250. Our approach relies on the analysis of community dissimilarity
392 indices, so that it can be applied to species-rich communities that contain a substantial amount of
rare species with low occurrence frequencies. The proposed approach is easy to conduct, since it
394 does not require any advanced statistical training. It allows performing a first exploratory analysis of
empirical data to assess the respective influences of complementary drivers of metacommunity
396 dynamics (Kingsolver & Schemske 1991, Shipley 2000).

Some ecological systems may deviate from our general predictions for a variety of reasons. Users
398 should then build alternative, biologically more relevant heuristic path models. Alternative path
models can include different set of paths between the variables of Fig. 2, they can incorporate
400 different signs for the predicted relationships, or they can incorporate other variables. For instance,
environmental variables may display cyclic temporal dynamics. In such cases, it may be more
402 pertinent to consider phase difference rather than absolute time difference (Δt). Similarly, the
implicit assumption of spatial isotropy in the path model that makes use of geographical distances Δx

may not be adapted to ecological systems in which the connectivity between sites is not solely driven by geographical distances (e.g., McRae et al. 2008). Another example is the one of disease or population outbreaks that travel through space and sometimes constitute a genuine environmental perturbation for entire communities (e.g., Tenow et al. 2013). In this case too, absolute time may not be a pertinent variable and may be fruitfully replaced by the state of outbreak (x-vt) where v is the speed of the travelling wave and x is the position of the site considered. More generally, when analysing transient systems, one should pay particular attention to the definition of paths involving time and to their interpretation. Our proposition is a simple and versatile approach to analyse standardized path coefficients, although this may not always be the choice to be favoured (Grace & Bollen 2005), so that researchers should evaluate the pros and cons of this choice for their particular case study.

Although the proposed framework appears powerful and robust, it is important to keep in mind that only simple linear relationships are modelled in the path analysis. Our analysis of simulated datasets supports this simple assumption (Fig. S4-90) and variable transformation procedures can be used to correct obvious non-linear relationships, as done here for some empirical case studies using log-transformation of geographical distances. Still, results should be solely interpreted as rough estimates of the respective influences of dispersal, demographic stochasticity and environmental filtering on community dynamics. Explored path models are not meant to be predictive. For such an endeavour, process-based dynamical models of metacommunity dynamics may be a much more suitable way forward (Evans et al. 2013, Mouquet et al. 2015). Such process-based dynamical models, however, require much more data on the system studied to be relevant. By identifying important drivers of metacommunity dynamics, the proposed framework can help design relevant process-based models that focus on the most influential processes.

The causal inference framework that we propose is a flexible way to make sense of temporal metacommunity data. We demonstrated its ability to detect influential ecological processes on simulated data. We illustrated its use on four real ecological datasets and explained how results can be interpreted. We finally explained how to deal with peculiarities of specific ecological systems within this framework, by modifying the path model. Beyond this methodological advance, our analyses of four sharply distinct ecological communities confirmed that environmental filtering is omnipresent and revealed that the environmental drivers of community composition do vary in both space and time, so that static metacommunity analyses should be abandoned in favour of a joint analysis of community turnover in both space and time.

Acknowledgments

This work has been funded by the Irstea-Carnot project “MetaRISC”. This research was financed by the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25). It has been partly undertaken at SETE, a laboratory which is part of the “Laboratoire d’Excellence” (LABEX) entitled TULIP (ANR-10-LABX-41). FM is funded by the CNRS and the ANR-funded ARSENIC and NGB projects (grants no. ANR-14-CE02-0012 and ANR-17-CE32-0011).

References

Anderson, M.J. & Cribble, N.A. (1998). Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Austr. Ecol.*, 23, 158-167.

Amarasekare, P. & Nisbet, R.M. (2001). Spatial heterogeneity, source-sink dynamics, and the local coexistence of competing species. *Am. Nat.*, 158, 572-584.

Arthaud, F., Vallod, D., Robin, J., & Bornette, G. (2012). Eutrophication and drought disturbance shape functional diversity and life-history traits of aquatic plants in shallow lakes. *Aquat. Sci.*, 74, 471-481.

Arthaud, F., Vallod, D., Wezel, W., Robin, J. & Bornette, G. (2013). Short-term succession of aquatic plant species richness along ecosystem productivity and dispersal gradients in shallow lakes. *J. Veg. Sci.*, 24, 148-156.

Azaele, S., Pigolotti, S., Banavar, J.R. & Maritan, A. (2006). Dynamical evolution of ecosystems. *Nature*, 444, 926-928.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134-143.

Beaumont, M.A. (2010). Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Sys.*, 41, 379-406.

Bertrand, C., Burel, F. & Baudry, J. (2016). Spatial and temporal heterogeneity of the crop mosaic influences carabid beetles in agricultural landscapes. *Land. Ecol.*, 31, 451-466.

Blanchet, S., Helmus, M.R., Brosse, S. & Grenouillet, G. (2014). Regional vs local drivers of phylogenetic and species diversity in stream fish communities. *Fresh. Biol.*, 59, 450-462.

Borcard, D., Legendre, P. & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, 73, 1045-1055.

Brokaw, N.V. (1985). Gap-phase regeneration in a tropical forest. *Ecology*, 66, 682-687.

Brown, B.L., Sokol, E.R., Skelton, J. & Tornwall, B. (2017). Making sense of metacommunities: dispelling the mythology of a metacommunity typology. *Oecologia*, 183, 643-652.

Buckley, H.L., Day, N.J., Case, B.S., Lear, G. & Ellison, A.M. (2018). Multivariate methods for testing hypotheses of temporal community dynamics. *bioRxiv*, 362822.

Calcagno, V., Mouquet, N., Jarne, P. & David, P. (2006). Coexistence in a metacommunity: the competition-colonization trade-off is not dead. *Ecol. Lett.*, 9, 897-907.

Chase, J.M. & Leibold, M.A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago, Ill.

Chave, J., & Leigh Jr, E. G. (2002). A spatially explicit neutral model of β-diversity in tropical forests. *Theor. Popul. Biol.*, 62, 153-168.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31, 343-366.

Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecol. Lett.*, 8, 1175-1182.

Datry, T., Larned, S.T., Fritz, K.M., Bogan, M.T., Wood, P.J., Meyer, E.I. *et al.* (2014). Broad-scale patterns of invertebrate richness and community composition in temporary rivers: effects of flow intermittence. *Ecography*, 37, 94-104.

Datry, T., Bonada, N. & Heino, J. (2016). Towards understanding the organisation of metacommunities in highly dynamic ecological systems. *Oikos*, 125, 149-159.

Evans, M.R., Bithell, M., Cornell, S.J., Dall, S.R., Díaz, S., Emmott, S. *et al.* (2013). Predictive systems ecology. *Proc. R. Soc. B*, 280, 20131452.

Fournier, B., Mouquet, N., Leibold, M.A. & Gravel, D. (2017). An integrative framework of coexistence mechanisms in competitive metacommunities. *Ecography*, 40, 630-641.

Fourtune, L., Paz-Vinas, I., Loot, G., Prunier, J. G., & Blanchet, S. (2016). Lessons from the fish: a multi-species analysis reveals common processes underlying similar species-genetic diversity correlations. *Freshwater Biology*, 61, 1830-1845.

Fourtune, L., Prunier, J. G., Paz-Vinas, I., Loot, G., Veyssière, C., & Blanchet, S. (2018). Inferring causalities in landscape genetics: An extension of Wright's causal modeling to distance matrices. *The American Naturalist*, 191, 491-508.

Gascuel, F., Laroche, F., Bonnet-Lebrun, A.-S. & Rodrigues, A.S.L. (2016). The effects of archipelago spatial structure on island diversity and endemism: predictions from a spatially-structured neutral model. *Evolution*, 70, 2657-2666.

Gilbert, B. & Bennett, J.R. (2010). Partitioning variation in ecological communities: do the numbers add up? *J. App. Ecol.*, 47, 1071-1082.

Grace, J. B., & Bollen, K. A. (2005). Interpreting the results from multiple regression and structural equation models. *The Bulletin of the Ecological Society of America*, 86, 283-295.

Gravel, D., Canham, C. D., Beaudet, M., & Messier, C. (2006). Reconciling niche and neutrality: the continuum hypothesis. *Ecol. Lett.*, 9, 399-409.

Guichard, F. (2017). Recent advances in metacommunities and meta-ecosystem theories. *F1000Research*, 6, 610.

Hastings, A. (1980). Disturbance, coexistence, history, and competition for space. *Theor. Popul. Biol.*, 18, 363-373.

Heino, J., Melo, A.S., Siqueira, T., Soininen, J., Valanko, S. & Bini, L.M. (2015). Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Fresh. Biol.*, 60, 845-869.

Henriques-Silva, R., Lindo, Z. & Peres-Neto, P.R. (2013). A community of metacommunities: exploring patterns in species distributions across large geographical areas. *Ecology*, 94, 627-639.

Hubbell, S.P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography,* Princeton University Press, Princeton, NJ.

Jabot, F. (2010). A stochastic dispersal-limited trait-based model of community dynamics. *J. Theor. Biol.* 262, 650-661.

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *Am. Nat.*, 104, 501-528.

Kalyuzhny, M., Kadmon, R. & Shnerb, N.M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecol. Lett.*, 18, 572-580.

Kingsolver, J. G. & Schemske, D. W. (1991). Path analyses of selection. *Trends Ecol. Evol.*, 6, 276-280.

Lamy, T., Jarne, P., Laroche, F., Pointier, J. P., Huth, G., Segard, A. *et al.* (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity. *Mol. Ecol.*, 22, 4445-4456.

Legendre, P. (2019). A temporal beta-diversity index to identify sites that have changed in exceptional ways in space–time surveys. *Ecol. Evol.*, 9, 3500-3514.

Legendre, P., & Gauthier, O. (2014). Statistical methods for temporal and space–time analysis of community composition data. *Proc. Roy. Soc. B.*, 281, 2013.2728.

Leibold, M.A. & Mikkelson, G.M. (2002). Coherence, species turnover, and boundary clumping: elements of meta-community structure. *Oikos*, 97, 237-250.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.*, 7, 601-613.

Leigh, E. G., Wright, S. J., Herre, E. A., & Putz, F. E. (1993). The decline of tree diversity on newly isolated tropical islands: a test of a null hypothesis and some implications. *Evol. Ecol.* 7, 76-102.

Logue, J.B., Mouquet, N., Peter, H., Hillebrand, H. & Metacommunity Working Group. (2011). Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol. Evol.*, 26, 482-491.

MacArthur, R.H. & Wilson, E.O. (1967). *The Theory of Island Biogeography,* Princeton University Press, Princeton, NJ.

Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422, 714.

Magurran, A. E., & Henderson, P. A. (2010). Temporal turnover and the maintenance of diversity in ecological assemblages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 3611-3620.

Malard, F., Uehlinger, U., Zah, R. & Tockner, K. (2006). Flood-pulse and riverscape dynamics in a braided glacial river. *Ecology*, 87, 704-716.

Massol, F., Gravel, D., Mouquet, N., Cadotte, M.W., Fukami, T. & Leibold, M.A. (2011). Linking community and ecosystem dynamics through spatial ecology. *Ecol. Lett.*, 14, 313-323.

May, R. M. (1973). *Stability and complexity in model ecosystems* (Vol. 6). Princeton university press.

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*(10), 2712-2724.

Molofsky, J., Bever, J. D., & Antonovics, J. (2001). Coexistence under positive frequency dependence. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268, 273-277.

Mouquet, N. & Loreau, M. (2003). Community patterns in source-sink metacommunities. *Am. Nat.*, 162, 544-557.

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D. *et al.* (2015). Predictive ecology in a changing world. *J. App. Ecol.*, 52, 1293-1310.

Muneepeerakul, R., Bertuzzo, E., Lynch, H.J., Fagan, W.F., Rinaldo, A. & Rodriguez-Iturbe, I. (2008). Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature*, 453, 220.

Münkemüller, T. & Gallien, L. (2015). VirtualCom: a simulation model for eco-evolutionary community assembly and invasion. *Meth. Ecol. Evol.*, 6, 735-743.

Munoz, F., Grenié, M., Denelle, P., Taudière, A., Laroche, F., Tucker, C. *et al.* (2018). ecolottery: Simulating and assessing community assembly with environmental filtering and neutral dynamics in R. *Meth. Ecol. Evol.*, 9, 693-703.

Nuvoloni, F.M., Feres, R.J.F. & Gilbert, B. (2016). Species turnover through time: colonization and extinction dynamics across metacommunities. *Am. Nat.*, 187, 786-796.

Overcast, I., Emerson, B.C. & Hickerson, M.J. (2019) An integrated model of population genetics and community ecology. *J. Biogeo.,* 46, 816-829.

Peres-Neto, P.R. & Legendre, P. (2010). Estimating and controlling for spatial structure in the study of ecological communities. *Glob. Ecol. Biogeogr.*, 19, 174–184.

Peres-Neto, P.R., Legendre, P., Dray, S. & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology*, 87, 2614-2625.

Poulet, N., Beaulaton, L. & Dembski, S. (2011). Time trends in fish populations in metropolitan France: insights from national monitoring data. *J. Fish Biol.*, 79, 1436-1452.

Preston, F.W. (1962). The canonical distribution of commonness and rarity: Part I. *Ecology*, 43, 185-215.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software*, 48, 1-36.

Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling*, 7, 206-218.

Shmida, A. & Wilson, M.V. (1985). Biological determinants of species diversity. *J. Biogeogr.*, 12, 1-20.

Shoemaker, L.G. & Melbourne, B.A. (2016). Linking metacommunity paradigms to spatial coexistence mechanisms. *Ecology*, 97, 2436-2446.

Skellam, J.G. (1952). Studies in Statistical Ecology: I. Spatial Pattern. *Biometrika*, 39, 346-362.

Soininen, J. (2010). Species turnover along abiotic and biotic gradients: patterns in space equal patterns in time? *BioScience*, 60, 433-439.

Sokol, E.R., Brown, B.L. & Barrett, J.E. (2017). A simulation-based approach to understand how metacommunity characteristics influence emergent biodiversity patterns. *Oikos*, 126, 723-737.

Srivastava, D.S. & Lawton, J.H. (1998). Why More Productive Sites Have More Species: An Experimental Test of Theory Using Tree-Hole Communities. *Am. Nat.*, 152, 510-529.

Tenow, O., Nilssen, A. C., Bylund, H., Pettersson, R., Battisti, A., Bohn, U., et al. (2013). Geometrid outbreak waves travel across Europe. *Journal of Animal Ecology*, 82, 84-95.

Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, 75, 2-16.

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I. & Thies, C. (2005). Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management. *Ecol. Lett.*, 8, 857-874.

Ulrich, W., Kryszewski, W., Sewerniak, P., Puchałka, R., Strona, G. & Gotelli, N. J. (2017). A comprehensive framework for the study of species co-occurrences, nestedness and turnover. *Oikos*, 126, 1607-1616.

Vannette, R. L., & Fukami, T. (2017). Dispersal enhances beta diversity in nectar microbes. *Ecol. Lett.*, 20, 901-910.

White, E. P., Ernest, S. M., Adler, P. B., Hurlbert, A. H., & Lyons, S. K. (2010). Integrating spatial and temporal approaches to understanding species richness. *Phil. Trans. Roy. Soc. B.*, 365, 3633-3643.

608 **Tab. 1. Ecological processes and environmental spatiotemporal variations included in the six simulated scenarios.**

	Ecological processes			Environmental variables	
Scenario	**Global dispersal**	**Neighbour dispersal**	**Environmental filtering**	**Spatial (e_1)**	**Temporal (e_2)**
1	+	-	-	-	-
2	-	+	-	-	-
3	+	-	+	+	-
4	+	-	+	-	+
5	+	-	+	+	+
6	-	+	+	+	+

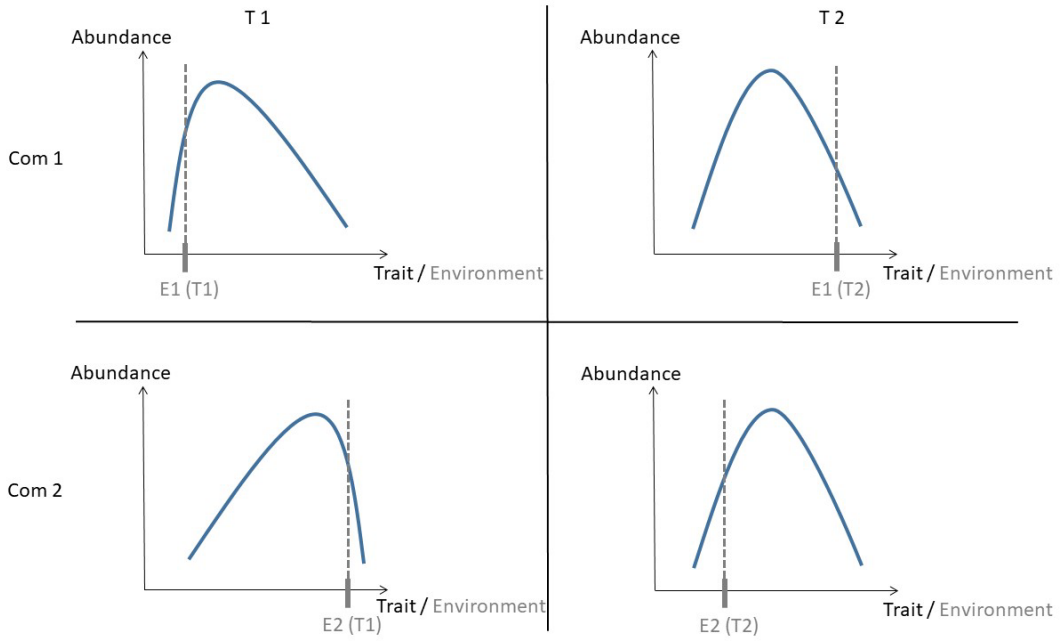
610

Box 1

Fig. 1. Thought experiment demonstrating the benefits of a spatiotemporal analysis of community composition.

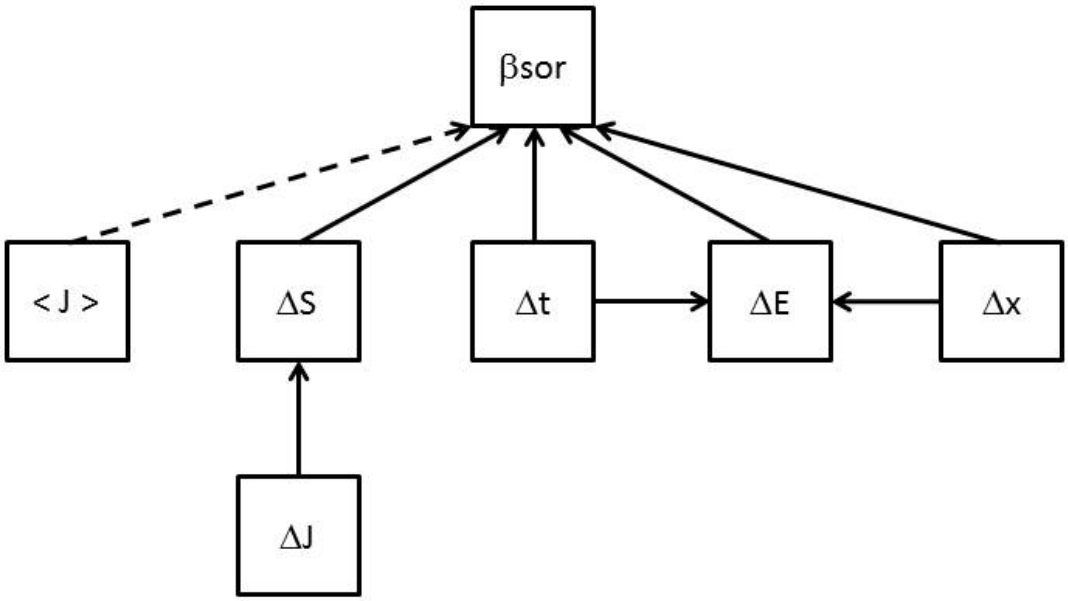
Imagine two communities with limited dispersal between each other and strongly structured by a one-dimensional environmental filter that selects species according to their (uni-dimensional) trait (Figure 1). At the first time step, community 1 experiences a low environmental value $E_1(t_1)$ and is mainly composed of species with low trait values (Fig. 1a), while community 2 experiences a large environmental value $E_2(t_1)$ and is mainly composed of species with large trait values (Fig. 1b). During the next time step, community 1 and 2 experiences large and low environmental values respectively, that enable some dispersers from the other community to settle in (Fig. 1c,d). This leads to a similar species composition of the two communities, despite the fact that current environmental conditions differ between the two communities. If one solely performs a spatial analysis of this metacommunity at time step 2, one would conclude that the environmental variable does not influence communities and thus that metacommunity dynamics is neutral with little dispersal limitation. If one instead looks at the temporal trajectories of the two communities (i.e. performs a spatiotemporal analysis), one would detect that community 1 and 2 moves to larger and lower trait values respectively, between t_1 and t_2, in a positively correlated manner with temporal environmental variations. One would also evidence that this tracking of environmental variation by community composition is only partial since abundant species do not have optimal trait values. This would provide evidence of some level of dispersal limitation between the two communities.

Fig. 1. Thought experiment demonstrating the benefits of a spatiotemporal analysis of community composition. Blue curves represent the trait distribution of the community.



638 **Fig. 2 Heuristic path model to test the signature of ecological processes on spatiotemporal diversity patterns.**

640 A dashed or plain arrow represents a negative or positive correlation, respectively. <J> stands for the average community size in the metacommunity, t for time, x for space, E for the local environment
642 and S for species richness. Δ values represent difference of statistics in space and time. For instance, because it controls the intensity of ecological drift, the average community size is expected to
644 negatively affect spatial and temporal diversity turnovers (negative arrow between <J> and β diversity).



646

Fig. 3. Path analyses on the six simulated scenarios.

Arrows depict significant effects. Arrow width represents the strength of the standardized estimates. Numerical values are reported in Table S3.

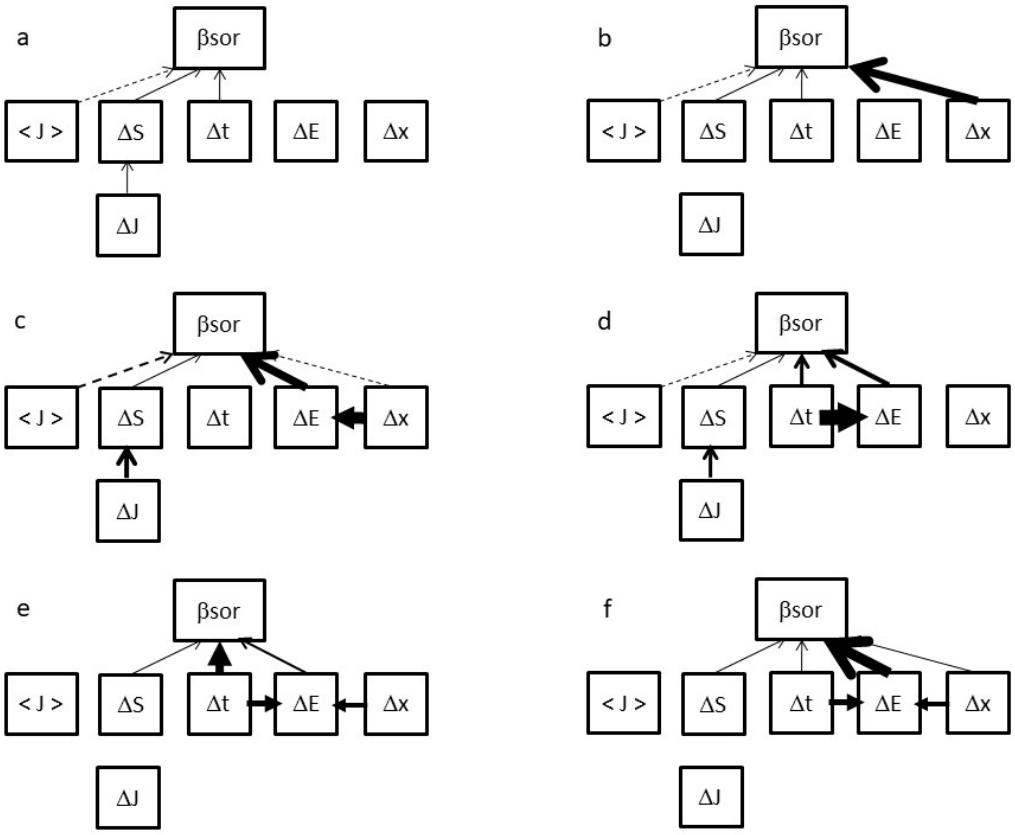
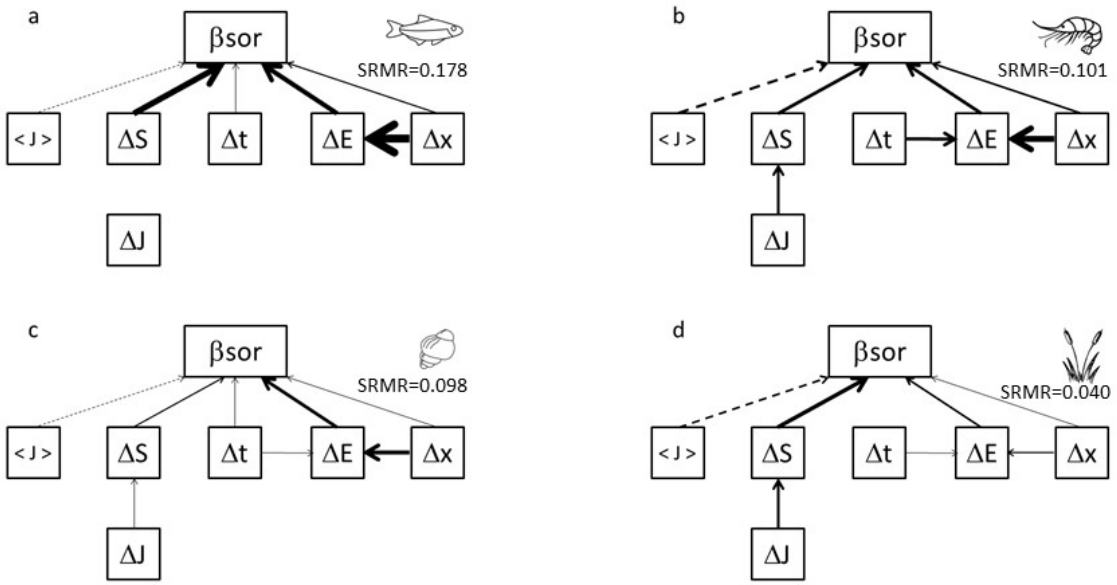


Fig. 4. Path analyses for the four empirical datasets.

a: freshwater fishes. b: aquatic invertebrates. c: molluscs. d: aquatic plants. Arrow width represents the strength of the standardized estimates. Dashed lines represent negative relationships. Paths from and towards ΔE were pooled in single arrows by summing the absolute values of the significant paths associated to each environmental variable. Only significant paths are shown. Numerical values of the standardized coefficients and of the associated p-values are reported in Tables S10-13. Values of the Standardized Root Mean Square Residual (SRMR) are mentioned for each dataset. Fish by Vladimir Belochkin, shrimp by Ana María Lora Macias, snail by Vega Asensio and cattail by Alex Muravev from the Noun Project.