



# Kent Academic Repository

Cai, Wang, Pichler, Maximilian, Biggs, Jeremy, Nicolet, Pascale, Ewald, Naomi, Griffiths, Richard A., Bush, Alex, Leibold, Mathew A., Hartig, Florian and Yu, Douglas W. (2025) *Assembly processes inferred from eDNA surveys of a pond metacommunity are consistent with known species ecologies*. *Ecography*, 2025 (6). ISSN 1600-0587.

## Downloaded from

<https://kar.kent.ac.uk/108626/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1111/ecog.07461>

## This document version

Publisher pdf

## DOI for this version

## Licence for this version

CC BY (Attribution)

## Additional information

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

## Research article

### Assembly processes inferred from eDNA surveys of a pond metacommunity are consistent with known species ecologies

Wang Cai<sup>1</sup>, Maximilian Pichler<sup>2</sup>, Jeremy Biggs<sup>3,4</sup>, Pascale Nicolet<sup>3</sup>, Naomi Ewald<sup>3</sup>, Richard A. Griffiths<sup>5</sup>, Alex Bush<sup>6</sup>, Mathew A. Leibold<sup>7,8</sup>, Florian Hartig<sup>2</sup> and Douglas W. Yu<sup>9,10</sup>✉

<sup>1</sup>CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, China

<sup>2</sup>Theoretical Ecology, University of Regensburg, Regensburg, Germany

<sup>3</sup>Freshwater Habitats Trust, Oxford, UK

<sup>4</sup>Oxford Brookes University, Oxford, OX3 0BP, UK

<sup>5</sup>Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, UK

<sup>6</sup>Lancaster Environment Centre, Lancaster University, UK

<sup>7</sup>Department of Biology, University of Florida, Gainesville, FL, USA

<sup>8</sup>Institut Natura e Teoria en Pirenèus, Surba, France

<sup>9</sup>Yunnan Key Laboratory of Biodiversity and Ecological Security of Gaoligong Mountain, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

<sup>10</sup>School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, UK

Correspondence: Douglas W. Yu ([dougwyu@mac.com](mailto:dougwyu@mac.com))

#### Ecography

2025: e07461

doi: [10.1111/ecog.07461](https://doi.org/10.1111/ecog.07461)

Subject Editor: Miguel G. Matias

Editor-in-Chief: Miguel Araújo

Accepted 08 January 2025



Technological advances are enabling ecologists to conduct large-scale and structured community surveys. However, it is unclear how best to extract information from these novel community data. We metabarcoded 48 vertebrate species from their eDNA in 320 ponds across England and applied the ‘internal structure’ approach, which uses joint species distribution models (JSDMs) to explain compositions as the result of four metacommunity processes: environmental filtering, dispersal, species interactions, and stochasticity. We confirm that environmental filtering plays an important role in community assembly, and find that species’ estimated environmental preferences are consistent with known ecologies. We also detect negative biotic covariances between fish and amphibians after controlling for divergent environmental preferences, consistent with predator–prey interactions (likely mediated by predator avoidance behaviour), and we detect high spatial autocorrelation for the palmate newt, consistent with its hypothesised relict distribution. Promisingly, ecologically and spatially distinctive sites are better explained by their environmental covariates and geographic locations, respectively, revealing sites where environmental filtering and dispersal limitation act more strongly. These results are consistent with the recent proposal that applying JSDMs to species distribution patterns can help reveal the relative importance of environmental filtering, dispersal limitation, and biotic interaction processes for individual sites and species. Our results also highlight the value of the modern interpretation of metacommunity ecology, which embraces the fact that assembly processes differ among species



[www.ecography.org](http://www.ecography.org)

© 2025 The Author(s). Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

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

and sites. We discuss how novel community data allow for several study design improvements that will strengthen the inference of metacommunity assembly processes from observational data.

Keywords: Aquatic eDNA, biodiversity, joint species distribution model (JSDM), macroecology, metabarcoding, *Triturus cristatus*

## Introduction

Metacommunity theory, which explicitly models feedback between local communities and regional species pools, has been proposed as a unifying theory of spatial community ecology (Leibold et al. 2004, Leibold and Chase 2018). In the framework of this theory, we consider a set of communities whose local population dynamics are governed by environmental filtering, species interactions, and ecological drift and that are additionally linked by dispersal. The goal of the theory is to understand how these four basic assembly processes determine species compositions in the metacommunity (Vellend 2016). Traditionally, this has been done with classical community data gathered by human observers, but the fact that modern sensors such as eDNA deliver community observations that are ideally suited for metacommunity analysis has created excitement in the field and also makes metacommunity analysis interesting for molecular ecologists (Hartig et al. 2024).

Empirical approaches to studying metacommunities mainly aim at inferring the relative contributions of the four assembly processes (dispersal, environmental filtering, species interactions, and drift) from empirical data. Examples of these are analyses based on community summary statistics, such as ordinations that describe different metacommunities using centroids and distances, and alpha and beta diversities (Fig. 1). Another common approach to analyse metacommunity data is variation partitioning, where, classically, community composition is explained by metacommunity-level contributions of environmental and spatial factors (Cottenie 2005). However, those approaches exhibit limited power to reveal assembly processes (Ovaskainen et al. 2019, Guzman et al. 2022), in part because summary metrics cannot reveal how the four processes differentially affect individual species and sites. Leibold et al. (2022) refer to such metacommunity-level metrics as studying the ‘external structure’ of metacommunities, because they assume that there is ‘one average mix’ of assembly processes that is the same across sites and species.

To avoid averaging assembly processes across sites and species, Leibold et al. (2022) propose studying the ‘internal structures’ of metacommunities, which dissect the importance of different assembly processes by each species and site. Technically, this can be done by using a joint species distribution model (JSDM) to partition the varying contributions of three model components (environmental covariates, species covariances, and spatial autocorrelation) to explain species presence/absence, for each individual species and site (Fig. 1). Among other things, this approach allows one

to relate environmental differences between sites, and trait differences between species, to differences in the variation explained by each component, thereby generating testable hypotheses that link distribution patterns to metacommunity assembly processes.

Simulation studies have shown that internal-structure analysis can indeed differentiate synthetic metacommunities that differ in site environmental distinctiveness and in species niche breadth, dispersal ability, niche centrality, and the presence or absence of competitive interactions (Fig. 1, Leibold et al. 2022, Terry et al. 2023). While these simulation results are encouraging, real metacommunity datasets have more complicated properties, including detection failures, measurement errors, and model uncertainty; not all species, environmental covariates, and sites can be included; species interact in multiple ways; and real metacommunities might be non-stationary, not least because of climate change (Abrego et al. 2021, Terry et al. 2023, Kadoya et al. 2024). Thus, it is important to gain more experience about the applicability of the internal structure idea to real data.

An ideal empirical metacommunity dataset for inferring internal structure would 1) consist of many local-community inventories with standardised species presence–absence or abundance information, 2) be within an area that is connected (and large) enough for dispersal (and dispersal limitation) to operate, 3) have the taxonomic breadth to include interacting guilds such as predators and prey, and 4) have measures of local environmental conditions relevant to the niche requirements of all these species. An exemplary study is provided by Kadoya et al. (2024) who applied internal-structure analysis to gillnet survey data covering three countries, 93 fish species, and 1853 lakes, and found that environmental covariates explained the most variation in species distributions and lake compositions, highlighting the importance of environmental filtering. Kadoya et al. (2024) then projected the effect of future climate heating on lake species compositions by running the fitted model with higher values of the degree-days environmental covariate while using the biotic covariances to simulate the effect of species interactions.

A promising alternative to traditional community observations is eDNA metabarcoding, which can generate repeated, large-scale, structured, and standardised community surveys (Hartig et al. 2024), but eDNA has so far rarely been used in metacommunity ecology (Vass et al. 2022, Macher et al. 2024).

Our survey data come from ponds in the south Midlands of England that were originally sampled to detect the great crested newt *Triturus cristatus*, a UK-protected amphibian species that breeds in ponds (Biggs et al. 2015).

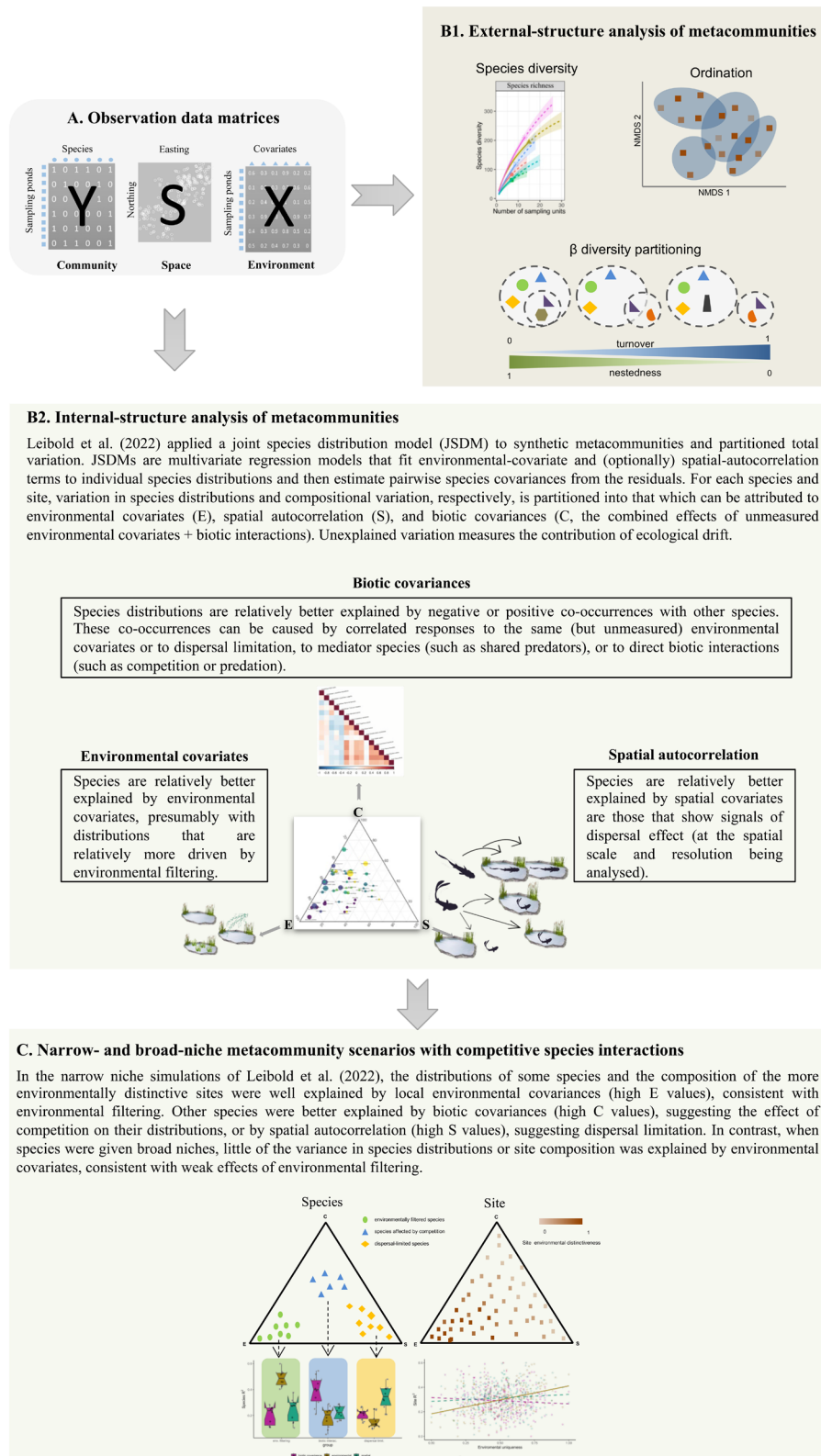


Figure 1. External- versus internal-structure analysis of metacommunities.

We metabarcode the residual eDNA to detect vertebrates, generating a community matrix of 320 ponds  $\times$  48 vertebrate species. Each pond was associated with eight environmental covariates and a geographic location, allowing us to fit three data matrices in a JSDM (Fig. 1). Here, we test how well the internal structure of a candidate pond metacommunity surveyed with eDNA matches expectations derived from external knowledge of species ecologies.

We expect to observe negative biotic covariances at the pond scale between newts and fishes. First, fish can reduce newt populations in two ways, via predation of eggs, larvae, and possibly adults; and via competition with newts for invertebrate prey (Beebee and Griffiths 2005, Hartel et al. 2007, Winandy et al. 2017). Second, Hartel et al. (2007) visually surveyed ponds and found negative correlations at the pond scale between predatory fish and two newt species, great crested *T. cristatus* and smooth *Lissotriton vulgaris*, but no correlation with common toad *Bufo bufo*, which is protected by bufotoxin (see also Hartel et al. 2007). Third, while this negative correlation could in theory be driven entirely by divergent environmental preferences along unmeasured niche axes, a behavioural mesocosm experiment with the alpine newt *Ichthyosaura alpestris* (Winandy et al. 2017) found that adult newts actively disperse away from aquaria with fish, but stay longer in otherwise equivalent aquaria without fish. Since newts are amphibious and can thus switch ponds more easily than fish can, this behaviour can rapidly generate negative residual covariances between newts and fish, except when ponds are isolated.

We also expect to see environmental covariate effects on newt distributions, with visual survey studies (Denoël and Lehmann 2006, Hartel et al. 2007, 2010, Denoël et al. 2013) finding greater newt occupancy in larger, deeper, more vegetated ponds near other ponds, especially when also inhabited by newts. Finally, we expected to obtain spatial covariance effects that are consistent with dispersal limitation, consistent with other freshwater community studies that have partitioned variation among biotic, environmental, and spatial components (Padial et al. 2014, Montaña et al. 2022).

## Material and methods

### Environmental covariates

To quantify land cover around each pond, we used Rowland et al.'s (2017) 21 UK land classes. For each pond, we calculated the proportions of land class within a 500 m radius of its point location and used principal component analysis in {FactoMineR} ver. 2.4 (Lê et al. 2008) to extract the top three principal components (accounting for 40% of total variation, Supporting information), which correlate with the degree of agriculture versus urban cover, grassland cover, and woodland cover. Each pond was also scored during sampling for ten standard pond variables used by surveyors to calculate the pond's habitat suitability index (HSI) for the great crested newt (ARG-UK 2010), of which we used five (Supporting information).

### Pond water sampling and metabarcoding assays

The pond water samples were the result of a single-season, great crested newt survey of 544 ponds in the south Midlands of England, UK, in 2017. Samples were collected and processed following Biggs et al. (2015) and were stored at ambient temperature until shipped to a commercial lab (NatureMetrics, Egham, UK), and DNA was extracted using a precipitation protocol (Tréguier et al. 2014), after which each sample's DNA was cleaned and subjected to 12 separate qPCR tests. After the qPCR assays, the residual eDNA was stored at  $-80^{\circ}\text{C}$ .

In 2019, the residual eDNA samples were retrieved and subjected to metabarcoding at NatureMetrics (PCR) and at Kunming Institute of Zoology (library preparation). The PCR and library preparation were conducted using a twin-tagging protocol (Yang et al. 2021). The resulting products were then sequenced on an Illumina HiSeq platform (PE150) at Novogene Tianjin, China. We processed the raw sequence data with the modified DAME bioinformatics pipeline of Cai et al. (2021). After sequence clustering, we generated a table of 540 ponds by 74 operational taxonomic units (OTUs). We assigned taxonomies to the OTUs using PROTAX (Somervuo et al. 2017, Axtner et al. 2019), setting prior probabilities to 0.90 for a list of expected UK vertebrate species (Harper et al. 2018). Further details regarding the metabarcoding process can be found in the Supporting information.

### Joint species distribution modelling

To fit JSDMs to the observed community data, we converted OTU read counts to presence-absence data. We retained only OTUs present in  $\geq$  five ponds and only sites with  $\geq$  one targeted OTU (= vertebrate species present in the UK), which reduced the number of OTUs from 74 to 48 and the number of ponds from 540 to 320. We assigned species-level taxonomies to OTUs that received  $\geq$  98% PROTAX probability of species assignment, and we classified the OTUs into six trait groups: fish, amphibians, perching birds, waterfowl, mammals, and domestic species. Domestic species are whose distributions we deemed as determined largely by humans (Supporting information).

We fit our data with two distinct JSDM structures. One model was fit to all the species in our dataset, terrestrial and aquatic (320 ponds  $\times$  48 species), and the other model fit to only the aquatic species (amphibians and fish, 279 ponds  $\times$  15 species; fewer ponds because we excluded those without aquatic species). A priori, aquatic species should be more likely to be filtered by pond characteristics, which make up five of our eight environmental covariates. Thus, we expect the aquatic species, especially the fish, to act more like narrow-niche species, and the terrestrial species to act more like broad-niche species (Fig. 1C).

All models were fitted using {sjSDM} ver. 1.0.6 (Pichler and Hartig 2021) running under R ver. 4.2.2 (www.r-project.org). We used a binomial likelihood and a multivariate probit link, linear main effects for the eight environmental



covariates, and a DNN (deep neural net) spatial model. To avoid overfitting, a light elastic net regularisation (Zou and Hastie 2005) was applied to all regression slopes and weights of the DNN (model fitting details in the Supporting information).

### Internal structure of the metacommunity

After fitting the two models (aquatic, terrestrial + aquatic), we used the ANOVA functions implemented in sjSDM (based on Leibold et al. 2022) to partition the variation of each species' distribution and each site's composition across environmental covariates (E), spatial autocorrelation (S), and biotic covariances (C) components. The relative partial McFadden  $R^2$  are visualised using ternary plots, where the positions of the species and sites reveal the relative contributions of E, S, and C: the metacommunity's internal structure (Fig. 1B2).

Leibold et al. (2022) found that environmentally more distinctive sites (at the ends of their one niche axis) received higher E values (implying a greater contribution of environmental filtering). To test for this result in a natural metacommunity, we regressed the individual pond partial E, S and C  $R^2$  values against pond environmental distinctiveness using quantile regression (50% quantile) (Fasiolo et al. 2021). We also tested the parallel hypothesis that geographically distinctive sites would have higher partial S  $R^2$  values (implying a greater contribution of dispersal limitation). In both cases, we defined distinctiveness as the leading eigenvector of the corresponding environmental or geographical euclidean distance matrix. We note that if predictor variables are collinear, bivariate correlations can be spurious and partial correlations calculated using multiple regressions should be preferred. However, environmental and geographic distinctiveness show practically no collinearity (Supporting information).

### Model generality

Our results are the outputs of a complex model that includes a linear environmental structure with eight environmental covariates and a DNN spatial structure with  $30 \times 2$  layers. Complex models run a risk of overfitting, so to estimate the risk of overfitting after elastic-net regularisation and to validate the predictive performance, we carried out a 20-fold cross-validation test with stratified multi-label sampling (Gunopulos et al. 2011, Szymański and Kajdanowicz 2017). The final explanatory and predictive area under the curves (AUCs) per species are the means over 20 folds. Species with higher predictive AUCs are those whose fitted models are more general (details in the Supporting information).

## Results

### Internal structure of the pond metacommunity

When analysing aquatic species only, we find that the fish are bimodally arrayed along the E–C axis, with four species relatively better explained by environmental covariates (higher

E values), and six species relatively better explained by biotic covariances (higher C values), which reflect the effects of unknown environmental covariates plus possible species interactions (Fig. 2A).

In contrast, none of the five amphibian species shows a high contribution of either environmental covariates or biotic covariances, but relative to fishes, amphibians show greater contributions of spatial effect (higher S values), especially the palmate newt *Lissotriton helveticus*, whose distribution is mostly explained by spatial effect. This species was detected in 16 ponds, in three separate sections of the survey area (Supporting information).

Including terrestrial species in the analysis (Fig. 2C) increases the relative contribution of biotic covariance for both fish and amphibians, which could reflect either the contributions of species interactions with terrestrial species or more unmeasured environmental covariates that have been revealed by adding the terrestrial species. The terrestrial species themselves also largely range along the E–C axis, with no clear clustering by trait group. Like the aquatic species, only one terrestrial species, Mandarin duck *Aix galericulata*, has a high S value (Supporting information) and is found in only five ponds. For the two site ternary plots (Fig. 2B, D), the general effect of adding terrestrial species is an increase in the variance accounted for by biotic covariances (site points shift upwards toward C).

We now examine species and site variation to try to infer some of the assembly processes that have resulted in these observed internal structures.

### Estimated environmental preferences

In the aquatic-only model (Fig. 3A), the pond effects for the fish species are in the direction of greater prevalence in larger ponds with lower risk of drying and less macrophyte cover. Several of the fish species are known to eat macrophytes, reduce macrophyte cover through other behaviours, and/or require higher oxygen with less macrophyte cover (Lopes et al. 2015, Maceda-Veiga et al. 2017, Stefanoudis et al. 2017). In contrast, for the amphibians, pond effects are in the direction of greater prevalence in smaller ponds with higher macrophyte cover. Pond drying risk, water quality, and shade showed essentially no effects on amphibian prevalence. Most of the effects of land cover on fish species are in the direction of lower prevalence in areas surrounded by agriculture or grassland. For amphibians, the effect is towards the prevalence of ponds bordered by woodland. This effect is observed for all five amphibian species but is only significant for the palmate newt and the common frog.

In the aquatic + terrestrial model (Fig. 3B), the effects of the environmental covariates on fish and amphibians remain largely the same as in the aquatic-only model. For the terrestrial species, most of the significant environmental-covariate effects are shade (% of pond perimeter shaded by trees), macrophyte cover, and land cover. Shade, which affects many perching birds and the grey squirrel, is most parsimoniously interpreted as increasing species detectabilities. Land-cover

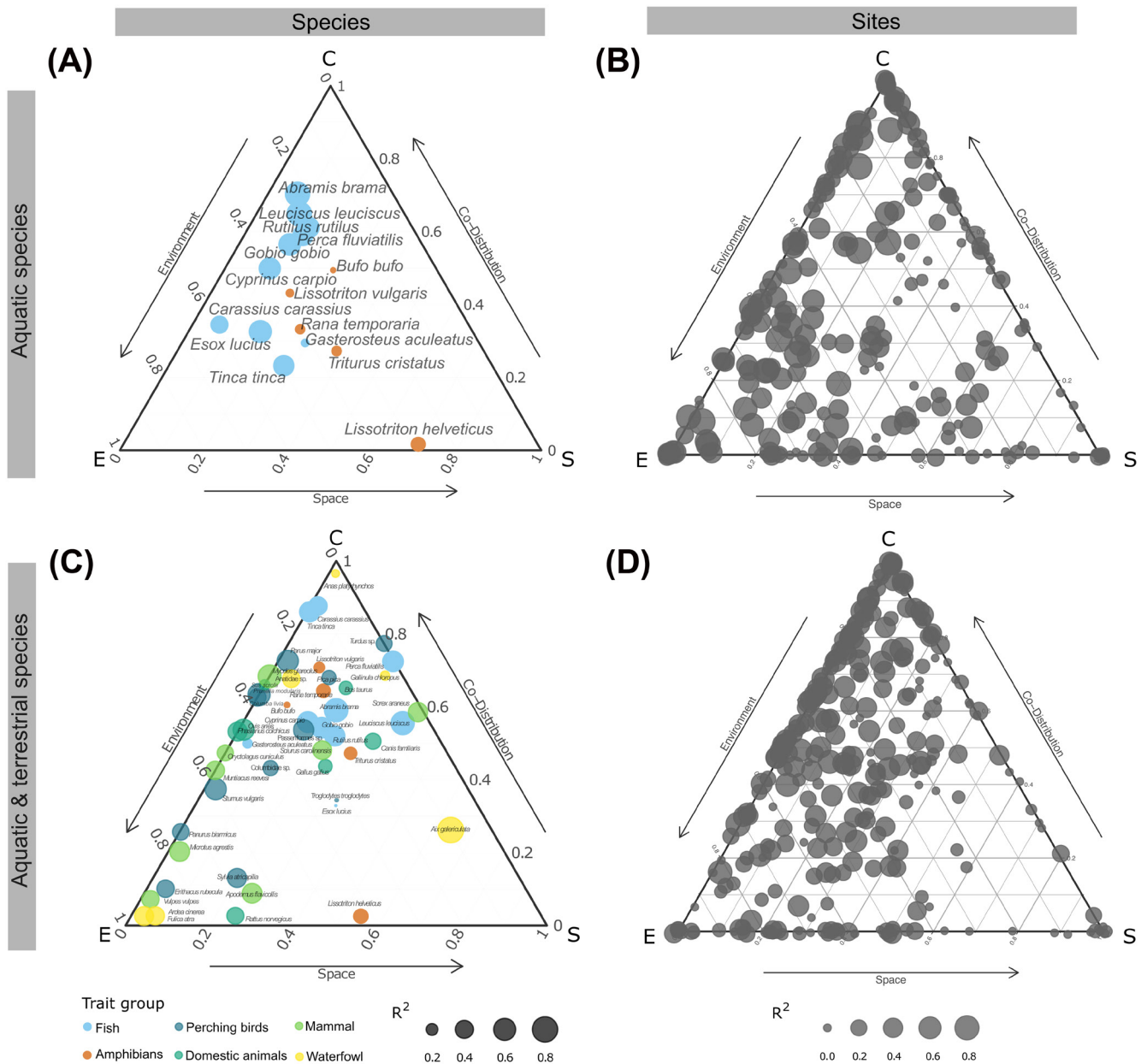


Figure 2. The internal structure of a pond metacommunity. The explained variation in species distributions or site compositions is decomposed, attributed to the three model components, environmental covariates (E), spatial autocorrelation (S), and biotic covariances (also known as co-distribution) (C), and visualised in a ternary plot after dividing each component's explained variance by its sum to allow comparison among species. Top row: aquatic species only. Bottom row: aquatic + terrestrial species. Left column: each point is a species, point size scales to total  $R^2_{\text{McFadden}}$  of each species, and the colours code for species trait group. Right column: each point is a site (pond), and point size scales to total  $R^2_{\text{McFadden}}$  of each site.

effects are variable across species, but we note that cows and sheep have higher prevalences in ponds bordered by ('improved') grassland.

### Biotic covariances

We visualise the residual biotic covariances in pairwise correlation plots (Fig. 4), where large absolute correlation-coefficient values correlate with high C-values in the internal-structure

ternary plots (Fig. 2 left column; linear model, aquatic species only,  $R^2=0.726$ ,  $p < 0.001$ ; aquatic + terrestrial,  $R^2=0.201$ ,  $p < 0.001$ ).

In the aquatic species model (Fig. 4A, B) and after filtering to the 2.5% most negative and positive values, the three surviving negative correlations are between the common frog *Rana temporaria* and two omnivorous fish species *Carassius carassius* and *Cyprinus carpio* and between great crested newt *Triturus cristatus* and a carnivorous fish *Esox lucius*. There are

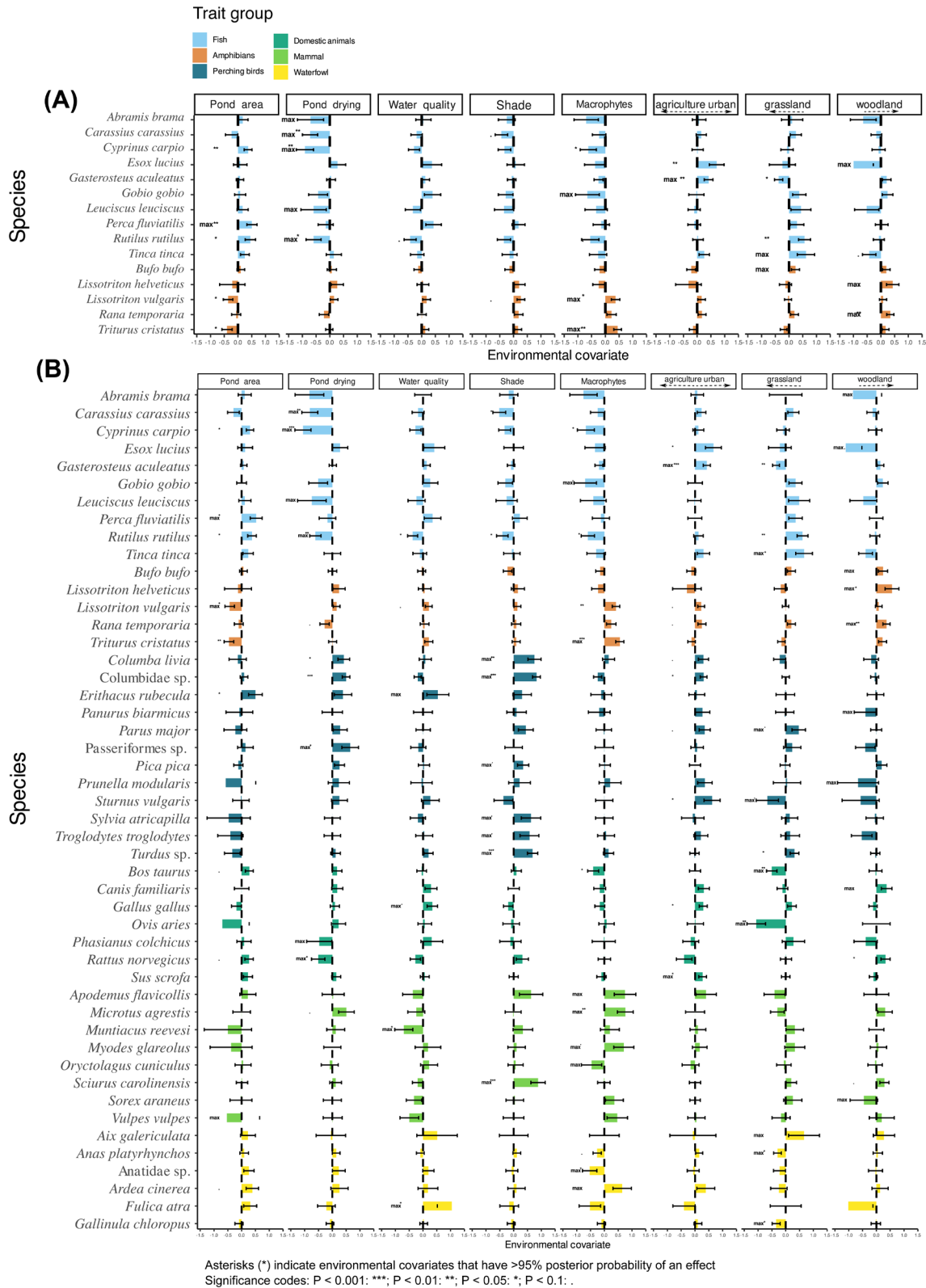


Figure 3. Estimated environmental effects. Eight environmental covariates were included in the model. The first five covariates from the left are taken from the ten standard pond variables used by surveyors to calculate the habitat suitability index (HSI) of each pond for the great crested newt and are therefore measured at all ponds in our dataset (ARG-UK 2010). The last three covariates describe the dominant land cover class within 500 m of each pond (more details in the Supporting information). Horizontal bars show the magnitudes, directions, and standard errors of the coefficients of each of the eight environmental covariates for each species. All covariates were normalised before fitting. Significance values are not corrected for multiple comparisons. Colours indicate species trait groups.



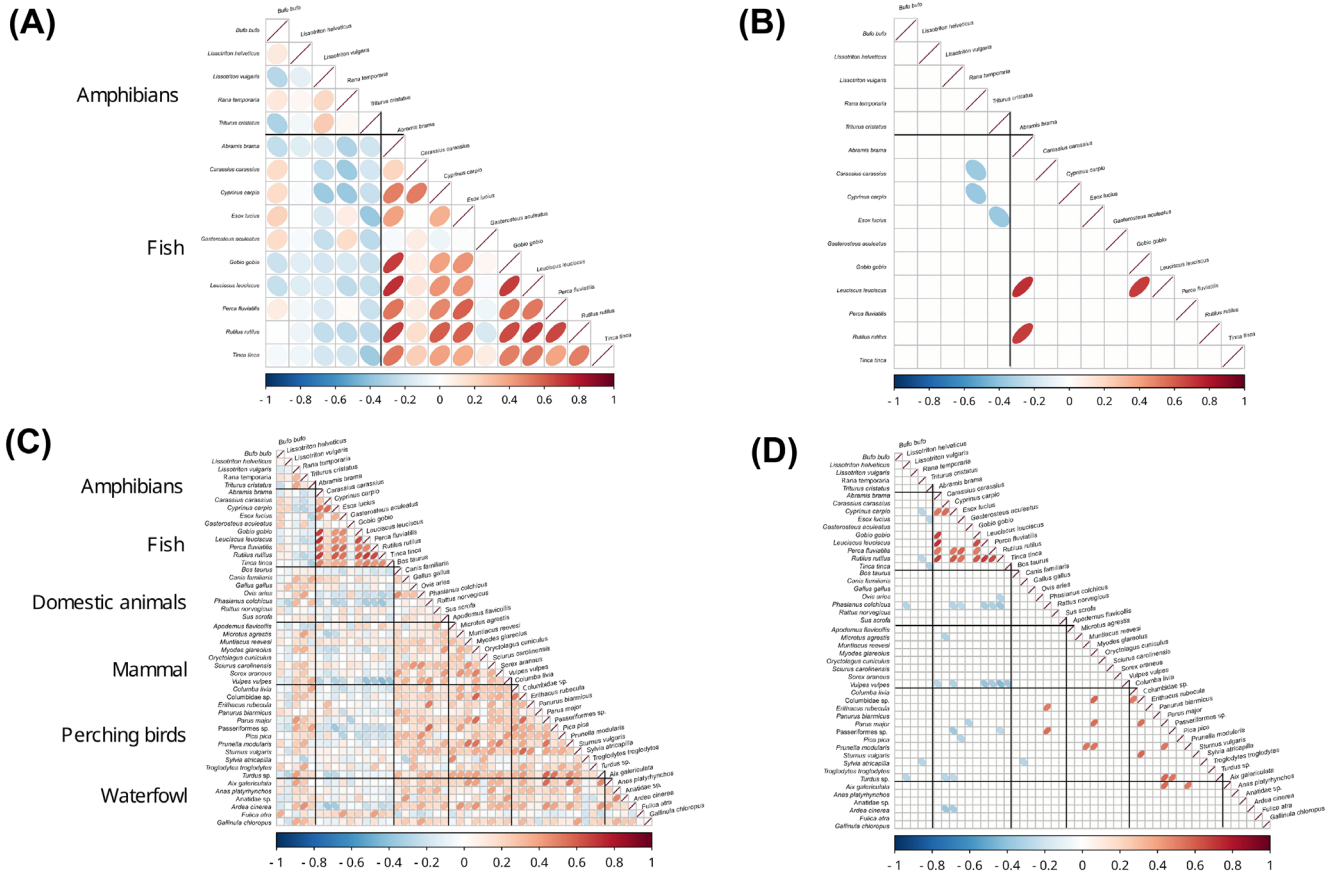


Figure 4. Biotic covariances. (A) Aquatic species only, all pairwise covariances. (B) Covariances filtered to the 2.5% most negative and positive. (C) Aquatic + terrestrial species, all pairwise biotic covariances. (D) Covariances filtered to the 2.5% most negative and positive.

also three surviving positive correlations between fish species, which we conservatively interpret as indicating unmeasured environmental covariates.

In the aquatic + terrestrial species model (Fig. 4C, D) and after filtering to the 2.5% most negative and positive values, there are four surviving negative correlations between amphibians and fish. The common frog is negatively correlated with two omnivorous fish species *Cyprinus carpio* and *Rutilus rutilus*, and great crested newt *Triturus cristatus* is negatively correlated with two carnivorous/omnivorous fish *Esox lucius* and *Tinca tinca*. Most of the surviving positive correlations occur among the fish species and among the bird species, which we again interpret as unmeasured environmental covariates. Also notable are negative correlations between several fish species with ring-necked pheasant *Phasianus colchicus* and red fox *Vulpes vulpes*.

## Relating distribution patterns to assembly processes

By site, the partial  $R^2$  explained by the environment increases significantly with the environmental distinctiveness of the site (Fig. 5A), and the partial site  $R^2$  explained by space increases significantly with the geographical distinctiveness of the site (Fig. 5B). This result holds up for four of the six

trait groups tested individually (amphibians, perching birds, domestic animals, and mammals) (Supporting information). Differing from Fig. 2, where the position of species in the internal structure depends on the relative weights of the three components (relative  $R^2$ ), the  $R^2$  here refers to the total contribution of a single component, regardless of its weight with other components. In other words, environmental filtering appears to be an increasingly more important assembly process for more environmentally distinctive sites, as predicted by Leibold et al. (2022), and spatial effect appears to be increasingly more important for geographically distinctive sites. Given this environment effect, we post hoc tested each covariate individually and found that the partial  $R^2$  explained by the environment increases only with pond area (Fig. 5C, Supporting information), suggesting that the species compositions of large ponds is determined more strongly by environmental filtering.

By species group, the partial  $R^2$  explained by biotic covariances is greatest for fish, amphibians, and waterfowl, and about equal with the  $R^2$  explained by the environment for the other three trait groups (Fig. 5D), which is consistent with species distributions being primarily governed by a combination of environmental filtering and (to a lesser extent) species interactions. The most pond-dependent species (fish,

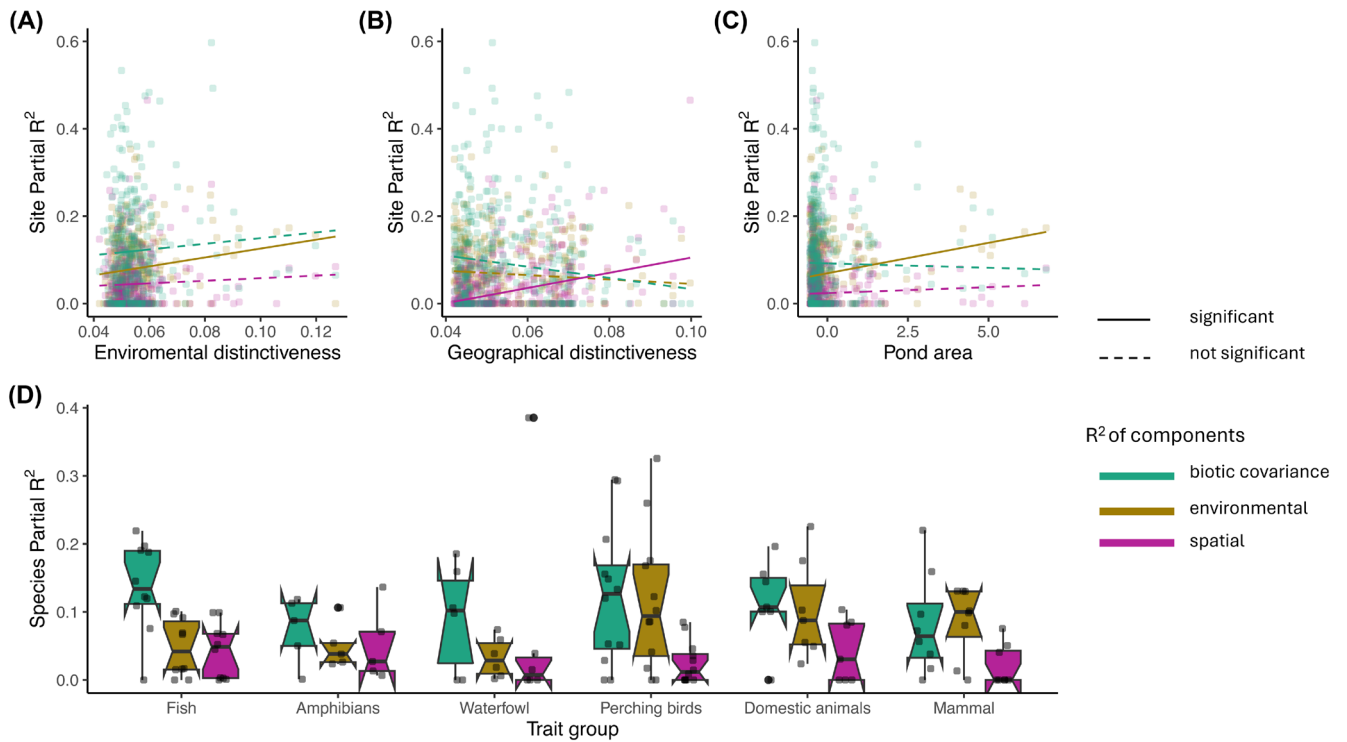


Figure 5. Correlation of the importance of assembly processes per site and species to environmental predictors and traits. (A)–(C) Quantile regression, correlating the importance of the three assembly mechanisms (measured by the share of absolute partial  $R^2$  values, each in a different colour) per site against (A) environmental distinctiveness ( $p < 0.001$  for the environmental line), (B) geographical distinctiveness ( $p < 0.001$  for the spatial line), and (C) pond area ( $p = 0.006$  for the environmental component line). (D) Association of  $R^2$  shares per species with traits, in this case, species groups. Note that the shared fractions were removed from the partial  $R^2$ , so the three components will not necessarily sum up to the total  $R^2$  value, which is displayed in Fig. 2.

amphibians, waterfowl) are the least well explained by our environmental covariates, although there are individual exceptions (Fig. 2A, C).

### Model generality

For the aquatic-only model, explanatory AUCs are always somewhat but not much greater than predictive AUCs, and explanatory and predictive AUCs are positively correlated (linear model, adjusted  $R^2 = 0.546$ ,  $p = 0.001$ ) (Fig. 6A). For the aquatic + terrestrial model, explanatory AUCs are again still always greater than predictive AUCs, but the correlation weakens considerably (linear model, adjusted  $R^2 = 0.139$ ,  $p = 0.005$ ), and for some terrestrial species, the model makes worse-than-random predictions (predictive AUCs  $< 0.5$ ) (Fig. 6B). The risk of overfitting is greater for low-predictive-AUC species, so the risk is greater for terrestrial species. We therefore focus on the aquatic species when interpreting model outputs.

## Discussion

The goal of our study was to reveal the internal structure of a real metacommunity, in order to infer the importances of different assembly processes per species and site. We estimated

the relative contributions of environmental covariates, biotic covariances, and space for explaining spatial variation in pond compositions. Pondscapes are convenient study systems because 1) each pond is unambiguously identified as a local community (De Meester et al. 2005), 2) there is an a priori division between pond niche aquatic versus terrestrial niche (Hill et al. 2021), 3) aquatic eDNA metabarcoding can efficiently generate hundreds of local-community inventories, and 4) the detected species encompass multiple trophic levels, increasing the possibility of detecting species interactions (Hering et al. 2018).

### Importance of taxonomic breath

We estimated two internal structures, one for aquatic species only and one for aquatic + terrestrial species (Fig. 2), which both showed that the distribution of each species and the species composition of each site were shaped by different mixes of ecological processes. This gives us our first conclusion, which is that the ‘one average mix’ approach to metacommunities indeed loses useful information contained in the variation among species and sites (Leibold et al. 2022). In this pond metacommunity, even if one makes the extreme assumption that all the variation partitioned to biotic covariance is also environmental filtering (but unmeasured), the species and sites still vary in their inferred degrees of dispersal

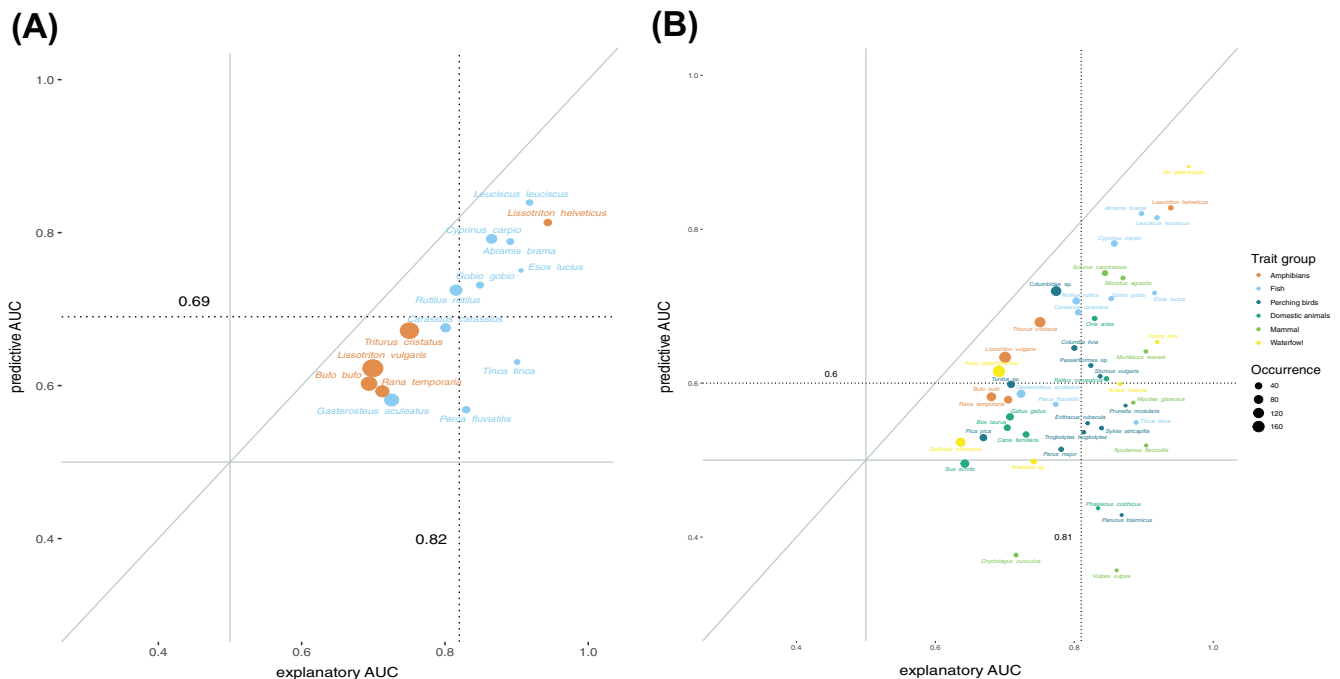


Figure 6. Predictive versus explanatory performance in two joint species distribution models (JSDMs). (A) Aquatic-species-only model. (B) Aquatic + terrestrial species model. Model performance was assessed using the AUC metric, and the dotted lines indicate mean AUC values (vertical for predictive and horizontal for explanatory). In both models, predictive performance is generally higher for fish than for amphibians, and explanatory performance is generally somewhat greater than the predictive performance, indicating moderate overfitting.

effect, and thus no single metacommunity paradigm (i.e. species-sorting, mass-effect, patch-dynamic, and neutral communities; Holyoak et al. 2005, Shoemaker and Melbourne 2016, Thompson et al. 2020, Suzuki and Economo 2021) can serve as an adequate description.

### Influence of environmental filtering

Looking at how metacommunity assembly processes vary across sites and species, both of the internal structures (Fig. 2A, C) suggest that environmental filtering is an important structuring force for many of the species in this pondscape. These findings are supported by the fact that environmental-covariate coefficient values are consistent with known biology. For instance, smooth and great crested newts are more prevalent in smaller ponds with greater macrophyte coverage (Fig. 3). In Romania, a visual survey found that macrophyte cover was also the strongest positive predictor of the great crested newt (Hartel et al. 2010), and in England, an eDNA pond survey also found that great crested newt is more prevalent in smaller ponds (Harper et al. 2020). Conversely, higher water quality is not associated with a higher prevalence of any amphibian species, despite its use as a component of the great crested newt's habitat suitability index (ARG-UK 2010). Indeed, Beebee and Griffiths (2005) and Sewell and Griffiths (2009) have argued that amphibians are not necessarily sensitive biological indicators of environmental quality.

Environmental filtering has also been implicated as the main determinant of macroinvertebrate species compositions

in pondscales in the UK Midlands (Hill et al. 2017) and the US south (Montaña et al. 2022). Heino et al. (2017) found that insects, macrophytes, and fish were better predicted by environmental filtering than by spatial effects in Finnish streams.

That said, observed niche preferences (environmental covariate values in fitted JSDMs) represent realised niches, not necessarily fundamental niches (Poggiato et al. 2021), and in the case of amphibians, parts of their realised niches are probably shaped by the predator-avoidance behaviour that we hypothesised in the Introduction drives the negative biotic covariances between newts, common frog, and fish (Fig. 4). As a result, the observed amphibian preferences for smaller ponds might disappear if fish were to disappear, since great crested and smooth newts have been observed to preferentially breed in larger ponds in continental Europe (Denoël and Lehmann 2006, Rannap and Briggs 2006, Skei et al. 2006, Denoël et al. 2013).

To the extent that environmental filtering is an important assembly process in this metacommunity, Leibold et al. (2022) predict that more environmentally distinctive sites should be more strongly determined by environmental filtering. This is what we indeed observe (Fig. 5A). Moreover, it appears that pond size is a major driver of uniqueness, since the compositions of large ponds are determined more strongly by environmental filtering (Fig. 5C). Tornero et al. (2024) censused macroinvertebrates in ponds and also found that environmentally more distinctive ponds, including larger ponds, are more compositionally distinctive.

## Influence on space on the community assembly

Looking at the contribution of space, one interpretation of which is a proxy for dispersal limitation in metacommunities (Leibold et al. 2022), the dataset does not show that different trait groups – as we defined them – differ consistently in the degree of spatial autocorrelation (Fig. 2C and 5D), contrary to our initial hypothesis.

Instead, we find that spatial effects are dominant for only two species, palmate newt and mandarin duck (Fig. 2). The spatial distribution of the 16 palmate newt detections is visibly patchy (Supporting information), which may reflect the persistence of relictual populations with a historic distribution associated with woodland (Beebee and Griffiths 2000). Whether the palmate newt is truly dispersal limited depends on whether nearby woodland is truly environmentally equivalent to historical woodland as the palmate newt perceives it. If not equivalent, then the patchy distribution is better interpreted as environmental filtering (here, habitat loss) leading to population decline and fragmentation. Alternatively, if the palmate newt requires continuous woodland to disperse through, which is not available in our study area, then dispersal limitation remains a viable hypothesis. Only individual tracking and experimental translocations and monitoring of population trajectories can answer this question definitively.

There are only five detections for mandarin duck (Supporting information), so we cannot conclude much about this species, but clearly a flighted bird should not be physically dispersal limited. Our working hypothesis is that this introduced species appears dispersal limited because it has successfully established in a range of pond environments, but the populations have not grown enough to start expanding. In short, a group of founder populations should appear dispersal limited.

Parallel to the environmental distinctiveness test (Fig. 5A), we found that geographically distinctive sites show stronger signals of dispersal limitation (Fig. 5B), or in other words, isolated ponds are better explained by their location than by their environmental conditions. Tornero et al. (2024) used network metrics to show that both active and passive dispersal macroinvertebrate communities in ponds are more compositionally unique the more isolated they are, which is suggestive of the same effect. More generally, pond studies report significant spatial effects (e.g. proximity to other ponds) on species compositions of both macroinvertebrates (Hill et al. 2017, Tornero et al. 2024) and vertebrates (Denoël and Lehmann 2006).

## Influence of co-distribution on community assembly

It is not a general rule that predators and prey should exhibit only negative biotic covariances, since predators search for prey, and they must overlap at some times and places, including over and above any shared environmental preferences. Theoretically, the covariances could go both directions, favouring a positive relationship when the prevalence of both species is equal, and possibly favouring a negative relationship

when the prevalence of one species is higher (Zurell et al. 2018). In fact, two remarkable studies that applied JSDMs to trawl data and to observational data from marine fisheries have reported both positive and negative biotic covariances between predators and prey (Astarloa et al. 2019, Zhang et al. 2022). However, Astarloa et al. (2019) found that most of the biotic covariances in their marine study system were negative, and they attribute this to predator-avoidance behaviour.

Keeping in mind the caveat that biotic (i.e. residual pairwise) covariances should not be taken as direct evidence for species interactions (Dormann et al. 2018, Zurell et al. 2018, Blanchet et al. 2020, Poggiato et al. 2021, Hartig et al. 2024), smooth newt, great crested newt, and common frog exhibited negative biotic covariances with nearly all the fish species (Fig. 4AC), and three (or four) of the covariances were among the 5% most extreme (Fig. 4BD). As suggested by Astarloa et al. (2019), these covariances are plausibly generated by the newt species actively avoiding predators (Winandy et al. 2017). The common frog is also known to avoid ovipositing in ponds that contain fish (or even ponds that were experimentally emptied of fish but still containing fish odour), even when alternative oviposition sites are pools at risk of drying (Kloskowski 2020, Kloskowski and Nieoczym 2022). In contrast, the bufotoxin-protected common toad showed both weak positive and negative covariances with fish (Fig. 4AC), and this species does not avoid ovipositing in fish-containing ponds (Kloskowski and Nieoczym 2022).

## What can eDNA bring to metacommunity ecology and internal structure analysis?

The large gains from eDNA metabarcoding in efficiency and error homogeneity over traditional survey methods make it feasible to generate datasets with many samples and many species, which can strengthen inference of metacommunity assembly processes. Most obviously, the large number of species detectable with eDNA increases the probability of detecting sets of species that are truly interacting, such as the negative correlations between fish and amphibians in this study (Fig. 4). Of particular interest would be adding data on pond macroinvertebrates and fungal diseases, which are obvious candidates for determining vertebrate distributions via species interactions (Beebee and Griffiths 2005).

Including more species could also suggest important but unmeasured environmental covariates via their pairwise species covariances. For example, while terrestrial and aquatic species mostly do not interact, some terrestrial species could be proxies for unmeasured land uses that affect aquatic species, such as increased agricultural runoff. In our models, we included agricultural and improved-grassland land cover types as environmental covariates, so we saw this effect directly (Fig. 3). The negative covariances of foxes and pheasants with multiple fish species (Fig. 4D) might be revealing other unmeasured land-use covariates. Specifically, > 35 million pheasants are released annually in the UK, mostly in England, and appear to boost fox numbers and incentivise land-cover management measures (Sage et al. 2020).



More technically, eDNA sampling makes it more feasible to collect multiple sample replicates, which would allow combining a JSDM with a detection model to account for observation error (Guillera-Arroita et al. 2017, Tobler et al. 2019, Doser et al. 2023, Diana et al. 2024, Hartig et al. 2024). Also, only two species in our dataset showed strong signals of dispersal limitation (Fig. 2), but this low number could be because near-neighbour ponds were not sampled in our dataset, removing the possibility of detecting fine-scale spatial autocorrelation and thereby possibly reducing the relative importance of dispersal that would support source–sink relations among closely adjacent ponds. Denser sampling might have detected more evidence of dispersal limitation. In our case, unfortunately, our dataset used the great crested newt sampling protocol, which requires only one sample per pond, and the ponds were dispersed across the landscape because the original use was to fit a species distribution model. Finally, eDNA sampling also makes it more feasible to survey ponds repeatedly over time, thereby creating a dataset that could be used to infer causality (Hartig et al. 2024). For example, if fish colonisation of ponds is consistently followed by loss of amphibians, we would have direct evidence supporting the causal hypothesis that we inferred from the negative biotic covariances between fish and amphibians.

### What can JSDMs and internal structure analysis bring to eDNA researchers?

Novel community datasets, including eDNA metabarcoding, are multivariate abundance datasets; that is, each species is a response variable, and there are many of them. Before computing power was widely available, such datasets were generally first reduced to tractable dissimilarity matrices before visualisation and analysis (e.g. NMDS and constrained ordination), but this approach may lose information and generate artefacts (Warton et al. 2012). However, for over a decade, it has been possible to analyse multivariate abundance data directly (Ovaskainen et al. 2017, Warton 2022). Here we have shown that JSDMs and variation partitioning allow the simultaneous analysis of environmental covariates, biotic covariances, and spatial autocorrelations, and the outputs can be visualised and interrogated in powerful ways (Ovaskainen et al. 2017, Popovic et al. 2019, Leibold et al. 2022, van der Veen et al. 2022, Warton 2022, Terry et al. 2023, Hartig et al. 2024).

### Conclusion

In conclusion, our study demonstrates that the combination of eDNA data and the analytical approach of exposing the ‘internal structure’ of metacommunities using JSDMs is a powerful tool for examining community assembly processes in structured landscapes. For the pond metacommunity in this study, we could reveal interesting patterns of how the importance of community assembly processes differs across sites and species, and relate those differences to spatial, environmental, and trait predictors. These results are consistent

with other empirical studies (Mehner et al. 2021, Vass et al. 2022, Kadoya et al. 2024) that support Leibold et al.’s (2022) inspiring insight that applying JSDMs to patterns of species distributions can help to reveal the relative importance of environmental filtering, dispersal limitation, and biotic interaction assembly processes on individual sites and individual species. Our study thus provides a blueprint for ecologists who want to study metacommunity processes and also for molecular ecologists who want to extract more information from their eDNA data.

*Acknowledgements* – We are grateful to the Freshwater Habitats Trust for collecting the samples and to NatureMetrics for donating the extracted DNA.

*Funding* – WC was supported by the Postdoctoral Fellowship of Xishuangbanna Tropical Botanical Garden, CAS, China Postdoctoral Science Foundation (2023M743747), Yunnan CAIYUN Postdoctoral Foundation and the Kunming Institute of Zoology. MAL is supported by NSF award 2025118. DWY was supported by the National Key Research and Development Program of China (2022YFC2602500), the Strategic Priority Research Program of Chinese Academy of Sciences, Grant (XDA20050202), the National Natural Science Foundation of China (41661144002), and the Yunnan Revitalization Talent Support Program: High-end Foreign Expert Project. This paper is an output of the sCom working group supported by sDiv, the Synthesis Centre of iDiv (DFG FZT 118, 202548816).

*Conflict of interest* – DWY is a cofounder of NatureMetrics, which provides commercial eDNA services. NatureMetrics is a minor shareholder in the NatureSpace Partnership. No further conflicts of interest to declare.

### Author contributions

**Wang Cai:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Resources (equal); Validation (equal); Visualization (equal); Writing - original draft (equal); Writing - review and editing (equal). **Maximilian Pichler:** Formal analysis (equal); Methodology (equal); Software (equal); Writing - review and editing (equal). **Jeremy Biggs:** Investigation (equal); Methodology (equal); Resources (equal); Writing - review and editing (equal). **Pascale Nicolet:** Investigation (equal). **Naomi Ewald:** Investigation (equal). **Richard A. Griffiths:** Resources (equal); Writing - review and editing (equal). **Alex Bush:** Methodology (equal); **Mathew A. Leibold:** Writing - review and editing (equal). **Florian Hartig:** Conceptualization (equal); Methodology (equal); Software (equal); Supervision (equal); Writing - review and editing (equal). **Douglas W. Yu:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing - original draft (equal); Writing - review and editing (equal).

### Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/eco.g.07461>.

## Data availability statement

Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.nk98sf7vr> (Cai et al. 2025).

## Supporting information

The Supporting information associated with this article is available with the online version.

## References

- Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N. M., Wang, J., Yu, D. W. and Ovaskainen, O. 2021. Accounting for species interactions is necessary for predicting how arctic arthropod communities respond to climate change. – *Ecography* 44: 885–896.
- ARG-UK 2010. ARG UK advice Note 5: Great crested newt habitat suitability index. – Amphibian and reptile groups of the United Kingdom <https://www.arguk.org/info-advice/advice-notes/9-great-crested-newt-habitat-suitability-index-arg-advice-note-5/file>.
- Astarloa, A., Louzao, M., Boyra, G., Martinez, U., Rubio, A., Irigoien, X., Hui, F. K. C. and Chust, G. 2019. Identifying main interactions in marine predator–prey networks of the Bay of Biscay. – *ICES J. Mar. Sci.* 76: 2247–2259.
- Axtner, J., Crampton-Platt, A., Horig, L. A., Mohamed, A., Xu, C. C. Y., Yu, D. W. and Wilting, A. 2019. An efficient and robust laboratory workflow and tetrapod database for larger scale environmental DNA studies. – *GigaScience* 8: giz029.
- Beebee, T. J. C. and Griffiths, R. A. 2000. Amphibians and reptiles: a natural history of the British Herpetofauna. The new naturalist. Vol. 1. – HarperCollins.
- Beebee, T. J. C. and Griffiths, R. A. 2005. The amphibian decline crisis: a watershed for conservation biology? – *Biol. Conserv.* 125: 271–285.
- Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., Foster, J., Wilkinson, J. W., Arnell, A., Brotherton, P., Williams, P. and Dunn, F. 2015. Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). – *Biol. Conserv.* 183: 19–28.
- Blanchet, F. G., Cazelles, K. and Gravel, D. 2020. Co-occurrence is not evidence of ecological interactions. – *Ecol. Lett.* 23: 1050–1063.
- Cai, W., Yang, C., Wang, X., Wu, C., Larrieu, L., Lopez-Vaamonde, C., Wen, Q. and Yu, D. W. 2021. The ecological impact of pest-induced tree dieback on insect biodiversity in Yunnan pine plantations, China. – *For. Ecol. Manage.* 491: 119173.
- Cai, W., Pichler, M., Biggs, J., Nicolet, P., Ewald, N., Griffiths, R. A., Bush, A., Leibold, M. A., Hartig, F. and Yu, D. W. 2025. Raw sequencing data for assembly processes inferred from eDNA surveys of a pond metacommunity are consistent with known species ecologies – Dryad Digital Repository, <https://doi.org/10.5061/dryad.nk98sf7vr>.
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. – *Ecol. Lett.* 8: 1175–1182.
- De Meester, L., Declerck, S., Stoks, R., Louette, G., Van De Meuter, F., De Bie, T., Michels, E. and Brendonck, L. 2005. Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. – *Aquat. Conserv.* 15: 715–725.
- Denoël, M. and Lehmann, A. 2006. Multi-scale effect of landscape processes and habitat quality on newt abundance: implications for conservation. – *Biol. Conserv.* 130: 495–504.
- Denoël, M., Perez, A., Cornet, Y. and Ficetola, G. F. 2013. Similar local and landscape processes affect both a common and a rare newt species. – *PLoS One* 8: e62727.
- Diana, A., Matechou, E., Griffin, J., Yu, D. W., Luo, M., Tosa, M., Bush, A. and Griffiths, R. A. 2024. eDNAPlus: a unifying modelling framework for DNA-based biodiversity monitoring. – *J. Am. Stat. Assoc.* 00: 1–23 <https://doi.org/10.1080/01621459.2024.2412362>.
- Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., Moretti, M. D., Pagel, J., Pinkert, S., Schleuning, M., Schmidt, S. I., Sheppard, C. S., Steinbauer, M. J., Zeuss, D. and Kraan, C. 2018. Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. – *Global Ecol. Biogeogr.* 27: 1004–1016.
- Doser, J. W., Finley, A. O. and Banerjee, S. 2023. Joint species distribution models with imperfect detection for high-dimensional spatial data. – *Ecology* 104: e4137.
- Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R. and Goude, Y. 2021. qgam: Bayesian nonparametric quantile regression modeling in R. *J. Stat. Soft.* 100: 1–31.
- Guillera-Aroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R. and Tingley, R. 2017. Dealing with false-positive and false-negative errors about species occurrence at multiple levels. – *Methods Ecol. Evol.* 8: 1081–1091.
- Gunopulos, D., Hofmann, T., Malerba, D. and Vazirgiannis, M. (eds) 2011. Machine learning and knowledge discovery in databases: European conference, ECML PKDD 2011, Athens, Greece, 5–9 September 2011, Proceedings, Part III. Lecture notes in computer science. – Springer.
- Guzman, L. M., Thompson, P. L., Viana, D. S., Vanschoenwinkel, B., Horváth, Z., Ptacnik, R., Jeliakzov, A., Gascón, S., Lemmens, P., Anton-Pardo, M., Langenheder, S., De Meester, L. and Chase, J. M. 2022. Accounting for temporal change in multiple biodiversity patterns improves the inference of meta-community processes. – *Ecology* 103: e3683.
- Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S. and Hänfling, B. 2018. Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). – *Ecol. Evol.* 8: 6330–6341.
- Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S. and Hänfling, B. 2020. Generating and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: a case study on a threatened amphibian. – *Environ. DNA* 2: 184–199.
- Hartel, T., Nemes, S., Cogălniceanu, D., Öllerer, K., Schweiger, O., Moga, C.-I. and Demeter, L. 2007. The effect of fish and aquatic habitat complexity on amphibians. – *Hydrobiologia* 583: 173–182.
- Hartel, T., Nemes, S., Öllerer, K., Cogălniceanu, D., Moga, C. and Arntzen, J. W. 2010. Using connectivity metrics and niche modelling to explore the occurrence of the northern crested newt *Triturus cristatus* (Amphibia, Caudata) in a traditionally managed landscape. – *Environ. Conserv.* 37: 195–200.
- Hartig, F., Abrego, N., Bush, A., Chase, J. M., Guillera-Aroita, G., Leibold, M. A., Ovaskainen, O., Pellissier, L., Pichler, M., Pog-

- giato, G., Pollock, L., Si-Moussi, S., Thuiller, W., Viana, D. S., Warton, D. I., Zurell, D. and Yu, D. W. 2024. Novel community data in ecology-properties and prospects. – *Trends Ecol. Evol.* 39: 280–293.
- Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. and Virtanen, R. 2017. Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. – *Oecologia* 183: 121–137.
- Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S., Hänfling, B., Kahlert, M., Leese, F., Meissner, K., Mergen, P., Reyjol, Y., Segurado, P., Vogler, A. and Kelly, M. 2018. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. – *Water Res.* 138: 192–205.
- Hill, M. J. et al. 2021. Pond ecology and conservation: research priorities and knowledge gaps. – *Ecosphere* 12: e03853.
- Hill, M. J., Heino, J., Thornhill, I., Ryves, D. B. and Wood, P. J. 2017. Effects of dispersal mode on the environmental and spatial correlates of nestedness and species turnover in pond communities. – *Oikos* 126: 1575–1585.
- Holyoak, M., Leibold, M. A. and Holt, R. D. 2005. Metacommunities spatial dynamics and ecological communities. – Univ. of Chicago Press.
- Kadoya, T., Nilsson, K. A., Kelly, J., Bartley, T. J., Säterberg, T., Guzzo, M. M., Esch, E., Koide, D., Matsuzaki, S. S., Terui, A., Akasaka, M. and MacDougall, A. S. 2024. Common processes drive metacommunity structure in freshwater fish. – *Global Ecol. Biogeogr.* 33: e13822.
- Kloskowski, J. 2020. Better desiccated than eaten by fish: distribution of anurans among habitats with different risks to offspring. – *Freshwater Biol.* 65: 2124–2134.
- Kloskowski, J. and Nieoczym, M. 2022. Strong behavioral effects of omnivorous fish on amphibian oviposition habitat selection: potential consequences for ecosystem shifts. – *Front. Ecol. Evol.* 10: 856258.
- Lê, S., Josse, J. and Husson, F. 2008. FactoMineR: an R package for multivariate analysis. – *J. Stat. Soft.* 25: 1–18.
- Leibold, M. A. and Chase, J. M. 2018. Metacommunity ecology. – Princeton Univ. Press.
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M. and Gonzalez, A. 2004. The metacommunity concept: a framework for multi-scale community ecology: the metacommunity concept. – *Ecol. Lett.* 7: 601–613.
- Leibold, M. A., Rudolph, F. J., Blanchet, F. G., De Meester, L., Gravel, D., Hartig, F., Peres-Neto, P., Shoemaker, L. and Chase, J. M. 2022. The internal structure of metacommunities. – *Oikos* 2022: e08618.
- Lopes, T. M., Cunha, E. R., Silva, J. C. B., Behrend, R. D. L. and Gomes, L. C. 2015. Dense macrophytes influence the horizontal distribution of fish in floodplain lakes. – *Environ. Biol. Fish.* 98: 1741–1755.
- Maceda-Veiga, A., López, R. and Green, A. J. 2017. Dramatic impact of alien carp *Cyprinus carpio* on globally threatened diving ducks and other waterbirds in Mediterranean shallow lakes. – *Biol. Conserv.* 212: 74–85.
- Macher, J.-N., Pichler, M., Creer, S., Martínez, A., Fontaneto, D. and Renema, W. 2024. Metacommunity theory and metabarcoding reveal the environmental, spatial, and biotic drivers of meiofaunal communities in sandy beaches. – *BioRxiv*. 2024.07.17.603914; doi: <https://doi.org/10.1101/2024.07.17.603914>.
- Mehner, T., Argillier, C., Hesthagen, T., Holmgren, K., Jeppesen, E., Kelly, F., Krause, T., Olin, M., Volta, P., Winfield, I. J. and Brucet, S. 2021. Model-based decomposition of environmental, spatial and species-interaction effects on the community structure of common fish species in 772 European lakes. – *Global Ecol. Biogeogr.* 30: 1558–1571.
- Montaña, C. G., Keppeler, F. W., Laughrey, C. P. and Schalk, C. M. 2022. Community assembly within ponds: the roles of space, time and environmental gradients. – *Aquat. Ecol.* 56: 123–142.
- Nemes, S., Demeter, L., Öllerer, K. and Hartel, T. 2008. Pond and landscape characteristics – which is more important for common toads (*Bufo bufo*)? A case study from central Romania. – *Appl. Herpetol.* 5: 1–12.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. and Abrego, N. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – *Ecol. Lett.* 20: 561–576.
- Ovaskainen, O., Rybicki, J. and Abrego, N. 2019. What can observational data reveal about metacommunity processes? – *Ecography* 42: 1877–1886.
- Padial, A. A., Ceschin, F., Declercq, S. A. J., De Meester, L., Bonecker, C. C., Lansac-Tôha, F. A., Rodrigues, L., Rodrigues, L. C., Train, S., Velho, L. F. M. and Bini, L. M. 2014. Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. – *PLoS One* 9: e111227.
- Pichler, M. and Hartig, F. 2021. A new joint species distribution model for faster and more accurate inference of species associations from big community data. – *Methods Ecol. Evol.* 12: 2159–2173.
- Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S. and Thuiller, W. 2021. On the interpretations of joint modeling in community ecology. – *Trends Ecol. Evol.* 36: 391–401.
- Popovic, G. C., Warton, D. I., Thomson, F. J., Hui, F. K. C. and Moles, A. T. 2019. Untangling direct species associations from indirect mediator species effects with graphical models. – *Methods Ecol. Evol.* 10: 1571–1583.
- Rannap, R. and Briggs, L. 2006. The characteristics of the great crested newt *Triturus cristatus* breeding ponds. Protection of the great crested newt. – Ministry of the Environment of the Republic of Estonia.
- Rowland, C. S., Morton, R. D., Carrasco, L., McShane, G., O’Neil, A. W. and Wood, C. M. 2017. Land cover map 2015 (vector, GB). – NERC Environmental Information Data Centre. <https://doi.org/10.5285/bb15e200-9349-403c-bda9-b430093807c7>
- Sage, R. B., Hoodless, A. N., Woodburn, M. I. A., Draycott, R. A. H., Madden, J. R. and Sotherton, N. W. 2020. Summary review and synthesis: effects on habitats and wildlife of the release and management of pheasants and red-legged partridges on UK lowland shoots. – *Wildl. Biol.* 2020: wlb.00766.
- Sewell, D. and Griffiths, R. A. 2009. Can a single amphibian species be a good biodiversity indicator? – *Diversity* 1: 102–117.
- Shoemaker, L. G. and Melbourne, B. A. 2016. Linking metacommunity paradigms to spatial coexistence mechanisms. – *Ecology* 97: 2436–2446.
- Skei, J., Ringsby, T., Rønning, L. and Dolmen, D. 2006. Habitat use during the aquatic phase of the newts *Triturus vulgaris* (L.) and *T. cristatus* (Laurenti) in central Norway: proposition for a conservation and monitoring area. – *Amphib. Reptilia* 27: 309–324.
- Somervuo, P., Yu, D. W., Xu, C. C. Y., Ji, Y., Hultman, J., Wirta, H. and Ovaskainen, O. 2017. Quantifying uncertainty of tax-

- onomic placement in DNA barcoding and metabarcoding. – *Methods Ecol. Evol.* 8: 398–407.
- Stefanoudis, P. V., Sayer, C. D., Greaves, H. M., Davidson, T. A., Robson, H., Almeida, D. and Smith, E. 2017. Consequences of fish for cladoceran, water beetle and macrophyte communities in a farmland pond landscape: implications for conservation. – *Fundam. Appl. Limnol.* 190: 141–156.
- Suzuki, Y. and Economo, E. P. 2021. From species sorting to mass effects: spatial network structure mediates the shift between metacommunity archetypes. – *Ecography* 44: 715–726.
- Szymański, P. and Kajdanowicz, T. 2017. A network perspective on stratification of multi-label data. – Presented at the Proceedings of Machine Learning Research. – *LIDTA* 74: 22–35.
- Terry, J. C. D., Langdon, W. and Rossberg, A. G. 2023. Codistribution as an indicator of whole metacommunity response to environmental change. – *Ecography* 2023: e06605.
- Thompson, P. L., Guzman, L. M., De Meester, L., Horváth, Z., Ptacnik, R., Vanschoenwinkel, B., Viana, D. S. and Chase, J. M. 2020. A process-based metacommunity framework linking local and regional scale community ecology. – *Ecol. Lett.* 23: 1314–1329.
- Tobler, M. W., Kéry, M., Hui, F. K. C., Guillera-Arroita, G., Knaus, P. and Sattler, T. 2019. Joint species distribution models with species correlations and imperfect detection. – *Ecology* 100: e02754.
- Tornero, I., Gascón, S., Cunillera-Montcusí, D., Sala, J., Compte, J., Quintana, X. D. and Boix, D. 2024. Untangling the roles of centrality and environmental contribution in diversity patterns across spatial scales. – *Ecosphere* 15: e4846.
- Tréguier, A., Paillisson, J.-M., Dejean, T., Valentini, A., Schlaepfer, M. A. and Roussel, J.-M. 2014. Environmental DNA surveillance for invertebrate species: advantages and technical limitations to detect invasive crayfish *Procambarus clarkii* in freshwater ponds. – *J. Appl. Ecol.* 51: 871–879.
- van der Veen, B., Hui, F. K. C., Hovstad, K. A. and O'Hara, R. B. 2022. Concurrent ordination: simultaneous unconstrained and constrained latent variable modelling. – *Methods Ecol. Evol.* 14: 683–695.
- Vass, M., Eriksson, K., Carlsson-Graner, U., Wikner, J. and Andersson, A. 2022. Co-occurrences enhance our understanding of aquatic fungal metacommunity assembly and reveal potential host–parasite interactions. – *FEMS Microbiol. Ecol.* 98: fiac120.
- Vellend, M. 2016. The theory of ecological communities. Monographs in population biology. – Princeton Univ. Press.
- Warton, D. I. 2022. Eco-Stats: data analysis in ecology: from t-tests to multivariate abundances. *Methods in statistical ecology*. – Springer.
- Warton, D. I., Wright, S. T. and Wang, Y. 2012. Distance-based multivariate analyses confound location and dispersion effects. – *Methods Ecol. Evol.* 3: 89–101.
- Winandy, L., Legrand, P. and Denoël, M. 2017. Habitat selection and reproduction of newts in networks of fish and fishless aquatic patches. – *Anim. Behav.* 123: 107–115.
- Yang, C., Bohmann, K., Wang, X., Cai, W., Wales, N., Ding, Z., Gopalakrishnan, S. and Yu, D. W. 2021. Biodiversity soup II: a bulk-sample metabarcoding pipeline emphasizing error reduction. – *Methods Ecol. Evol.* 12: 1252–1264.
- Zhang, Y., Zhang, C., Xu, B., Ji, Y., Ren, Y. and Xue, Y. 2022. Impacts of trophic interactions on the prediction of spatio-temporal distribution of mid-trophic level fishes. – *Ecol. Indic.* 138: 108826.
- Zou, H. and Hastie, T. 2005. Regularization and variable selection via the elastic net. – *J. R. Stat. Soc. B* 67: 301–320.
- Zurell, D., Pollock, L. J. and Thuiller, W. 2018. Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? – *Ecography* 41: 1812–1819.