



# Importance des processus de dispersion neutre dans la structuration spatiale des communautés de diatomées benthiques des cours d'eau, à l'échelle intercontinentale Implications en termes de biodiversité et de tolérance.

Thibault Leboucher

## ► To cite this version:

Thibault Leboucher. Importance des processus de dispersion neutre dans la structuration spatiale des communautés de diatomées benthiques des cours d'eau, à l'échelle intercontinentale Implications en termes de biodiversité et de tolérance.. Sciences de l'environnement. Thèse Université de Bordeaux, Ecole doctorale sciences et environnements ED 304, spécialité écologie évolutive, fonctionnelle et des communautés, 2018. Français. NNT : . tel-02608302

**HAL Id: tel-02608302**

**<https://hal.inrae.fr/tel-02608302>**

Submitted on 16 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE  
POUR OBTENIR LE GRADE DE  
**DOCTEUR DE  
L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS – E.D. 304  
SPÉCIALITÉ : ÉCOLOGIE ÉVOLUTIVE, FONCTIONNELLE ET DES COMMUNAUTÉS

Par **Thibault LEBOUCHER**

**Importance des processus de dispersion dans la structuration spatiale  
des communautés de diatomées benthiques des cours d'eau, à  
l'échelle intercontinentale.**

**Implications en termes de biodiversité et de tolérance.**

Sous la direction de Juliette ROSEBERY et de Sophia I. PASSY

Soutenue publiquement le 14 décembre 2018

Membres du Jury :

Mme **PASSY** Sophia I., Professeure associée, Université du Texas à Arlington, **co-directrice**  
M. **PRYGIEL** Jean, Professeur, Université de Lille, **rapporteur - Rapporteur de soutenance**  
Mme **ROSEBERY** Juliette, Chargée de recherche, Irstea Bordeaux, **directrice**  
M. **SABATER** Sergi, Professeur, Université de Gérone, **rapporteur - Président du Jury**  
M. **TABACCHI** Eric, Chargé de recherche, CNRS Toulouse, **rapporteur**  
M. **VYVERMAN** Wim, Professeur, Université de Gand, **examineur**



## **Importance des processus de dispersion dans la structuration spatiale des communautés de diatomées benthiques des cours d'eau, à l'échelle intercontinentale. Implications en termes de biodiversité et de tolérance.**

**Résumé :** La structuration des communautés biotiques apparaît aujourd'hui comme étant le résultat d'une interaction entre processus déterministes et stochastiques. Le *species sorting*, c'est-à-dire le couplage de l'influence des filtres environnementaux et des interactions biotiques, ainsi que les processus de dispersion apparaissent comme des facteurs déterminants pour la structuration des communautés, mais leur importance relative est encore mal comprise car conditionnée par les échelles spatiales, et les niveaux d'anthropisation considérés. Si des approches quantitatives multi-échelles sont donc vivement recommandées pour mieux comprendre comment les processus locaux et régionaux affectent les communautés, celles-ci sont encore trop rares dans la littérature. Nous avons ainsi tenté de mettre en place ce type d'étude, en prenant comme organisme modèle les diatomées, producteurs primaires jouant un rôle clef au sein des écosystèmes aquatiques. L'objectif était d'appréhender le rôle des processus déterministes et stochastiques dans la structuration des méta-communautés de diatomées benthiques en cours d'eau soumises à différents niveaux d'eutrophisation, de l'échelle locale à l'échelle intercontinentale. A l'aide de bases de données françaises et américaines comprenant plus de 4 000 stations d'échantillonnage, ce travail a permis de mettre en évidence une forte influence de l'eutrophisation tout au long du gradient spatial d'observation, induisant une baisse de la diversité  $\beta$ . La structuration des communautés diatomiques est de plus apparue fortement influencée par les processus déterministes locaux (*species sorting*) aux faibles niveaux d'anthropisation, puis par les processus de dispersion lorsque le niveau trophique augmentait. La dispersion limitée et l'effet de masse sont les deux processus évoqués pour expliquer cette forte influence de la dispersion sur la structuration des communautés. L'effet de masse permet à une espèce de subsister en dehors de son optimum environnemental grâce à l'arrivée constante d'individus via un processus d'immigration, induisant de fait une forte homogénéisation des communautés. Si la dispersion limitée est plutôt bien appréhendée aujourd'hui dans la littérature, l'effet de masse restait jusqu'ici difficile à mettre en évidence, constituant un verrou scientifique important à lever. Nos travaux ont permis de développer une méthode innovante capable de mesurer l'influence de ce processus sur les communautés de disperseurs passifs que sont les diatomées. Une meilleure compréhension de ce mécanisme représente une avancée conséquente, notamment dans le cadre du déclin mondial de la biodiversité. Ce phénomène impacte également significativement les méthodes de bio-indication de la qualité des milieux puisque la présence d'espèces dans des habitats qui leur sont *a priori* peu favorables peut potentiellement masquer l'effet des conditions locales du milieu sur les communautés, et ainsi induire un biais dans l'évaluation de l'état écologique des cours d'eau. Nos résultats vont dans ce sens, en révélant que la prise en compte de ce processus dans le calcul de l'Indice Biologique Diatomées permet une réponse plus fine de l'indice à l'environnement local, mettant ainsi en évidence l'importance de prendre en compte les processus de dispersion lors de l'élaboration de nouvelles méthodes de bio-indication.

**Mots-clefs :** Echelle intercontinentale ; Dispersion ; Patrons écologiques ; Effet de masse ; Assemblage des communautés



# **The importance of dispersion processes in the spatial structure of benthic diatom communities in rivers, at an intercontinental scale. Implications for biodiversity and vulnerability.**

**Abstract :** The structuration of biotic communities is generally caused by an interplay between stochastic and deterministic processes. Species sorting, i.e. environmental filtering coupled with biological interactions, and dispersal processes are particularly important drivers of communities, however, their relative influence are constrained by the effects of spatial scale and therefore poorly understood. Quantitative multi-scales approaches are thus recommended to better understand how local and regional processes affect communities. However, few studies have actually implemented quantitative scale-explicit frameworks to explore such relationships. This is one of our aims here, to quantify the relative contributions of local community assembly processes vs. stochastic processes in driving benthic diatom communities, and to assess how spatial scale and anthropogenic disturbance interact to dictate which driver dominates. Using French and US databases encompassing more than 4 000 spatially distinct localities, this work provides strong evidences stating that nutrient supply mediated community structuring whatever the scale considered. Nutrient enrichment was associated with an overall decline in diatom  $\beta$ -diversity and a shift in assembly processes from species sorting to dispersal. Dispersal limitation and mass effect are the two major dispersal processes acting on diatom communities. While dispersal limitation can be detected by experimental or modeling techniques, mass effect is more challenging to detect as it can allow species to persist in unfavorable habitats, i.e. far away from their environmental optimum. Here, we developed an innovative and robust method for the detection of this process for passive dispersers, such as diatoms. Beyond its importance in the taxonomic homogenization, causing extensive impairment of freshwater ecosystems, mass effect is also evoked for its potential impact on bio-assessment methods. Our results largely supported the idea that mass effect may induce a biased evaluation of water quality, since taking into account this process for the calculation of the Biological Diatom Index leads to a more acute assessment of river ecological status. We then advocate for a better integration of spatial mechanisms in any diatom-based indices, and by extension in any biotic index based on passive dispersers.

**Keywords :** Intercontinental scale ; Dispersion ; Ecological patterns ; Mass effect ; Community assembly

---

## **Unité de recherche**

Unité de Recherche Écosystèmes aquatiques et changements globaux (EABX)  
Irstea – 50 Avenue de Verdun, 33612 Cestas Cedex, France.



## Remerciements

Cette thèse a été financée par Irstea et l'Agence Française pour la Biodiversité. Merci à ces deux organismes pour cela, ainsi qu'au LabEx COTE qui a en partie financé trois mois de mobilité à l'Université du Texas à Arlington dans le laboratoire de Sophia Passy.

Je tiens à remercier Jean Prygiel, professeur à l'Université de Lille, Sergi Sabater, professeur à l'Université de Gérone (Espagne), ainsi qu'Eric Tabacchi, chercheur au CNRS (à Toulouse) pour avoir accepté d'être rapporteurs de cette thèse. Je tiens également à remercier Wim Vyverman, professeur à l'Université de Gand (Belgique), pour avoir accepté de faire partie du jury de soutenance.

Par où commencer, et comment résumer trois années qui ont changé tant de choses dans ma vie ? Ces trois années que j'ai pu passer en compagnie de si excellents et si admirables collègues, dont certaines et certains sont aujourd'hui des amies et amis proches que j'ai beaucoup de peine à quitter. La perte de mémoire inhérente à la rédaction d'un manuscrit de thèse, temporaire je l'espère, a ça de bon qu'elle permet d'oublier les mauvais souvenirs pour se concentrer sur l'essentiel... ou presque !

Juliette, je ne sais comment te dire à quel point j'ai été heureux de faire cette thèse sous ta direction. Je te remercie pour tout ce que tu as fait pour moi, pour ta gentillesse, ta bienveillance, ton humour et tous les moments passés à discuter. Dans le monde actuel de la recherche, toujours plus individualiste et obnubilé par une logique de chiffres et d'indicateurs, je n'ai pu qu'apprécier tes qualités et ton émerveillement pour l'étude du vivant. Je te dois énormément. Merci pour ta confiance, ta patience, tes conseils et ta disponibilité. Tu m'as aidé à grandir et j'espère sincèrement que j'aurai l'occasion de pouvoir continuer à travailler avec toi.

Sophia, thank you for co-supervising my thesis with Juliette and for welcoming me to your laboratory for these few months. I learned a lot during that time and I would have liked to stay longer in Arlington. It has been a chance and a pleasure to work with you and I hope we will have the opportunity to work together again.

Merci à toute l'équipe Carma/Ecovea pour votre bonne humeur et tous les bons moments partagés !! J'espère qu'en ces temps difficiles vous saurez conserver ce cadre de travail presque parfait qui donne envie de se lever le matin pour passer un peu de temps en bonne compagnie, au détour d'un café, d'un gâteau, d'une bière aux caves (après le boulot évidemment !) ou d'une sortie VTT... Merci pour toutes les discussions qui m'ont permis de faire avancer mes pensées et mon travail de recherche, mais aussi, ou peut être surtout, pour toutes les autres ! Les petits bonheurs de la vie tiennent parfois à peu de chose, et consciemment ou non, vous avez grandement participé aux miens ces trois dernières années. J'espère que nous aurons l'occasion de nous croiser à nouveau. Par contre Ecovea, c'est vraiment pas sérieux comme nom... Je me demande bien

qui a pu proposer le nom d'un système de douche à économie d'eau pour renommer l'équipe...

Au delà de l'équipe, je voudrais également transmettre quelques remerciements aux différentes personnes du centre qui ont à un moment ou un autre contribué au bon fonctionnement de cette thèse, aux secrétaires, gestionnaires, documentalistes, informaticiens. Merci à vous tous de nous permettre de mener à bien notre travail de recherche dans les meilleures conditions.

Comme je l'ai écrit plus haut, certaines et certains d'entre vous sont devenus des amies et amis proches. Merci à vous pour tous ces bons moments passés, ces fous rires, ces soirées en votre compagnie, ces week-ends portes ouvertes, ces parties de Carcassonne, Colt Express, Imagine et autres jeux, ces week-ends filles/mixte... Tous ces bons moments qui en appellent pleins d'autres je l'espère... Merci à vous d'avoir été là trois ans durant et d'avoir bien souvent été mes repères. Merci pour l'aide que vous avez aussi pu m'apporter pendant la rédaction, en débloquent mes problèmes avec Latex, avec ggplot, ou encore pour ne m'avoir pas (trop) tenu rigueur de mes moments d'absence pendant les pauses... J'ai une pensée particulière pour une personne qui se reconnaîtra... Merci pour ton soutien, ton et tes attention(s) au quotidien et en particulier dans les moments plus difficiles.

Merci à vous, Papa, Maman et Beno, merci d'avoir toujours été là pour moi. Merci de m'avoir donné goût à la nature, à la randonnée, à la musique, à la littérature... Toutes ces petites choses qui font de moi qui je suis aujourd'hui. Merci aussi d'avoir pris soin de mon chaton en particulier pendant cette période de rédaction, qui, malgré tout l'amour que je lui porte, aurait peut être fini en orbite...

Enfin merci à toi, Julie, merci d'être là, merci pour ton soutien, merci pour tes petites attentions, merci pour tout... Je ne sais pas comment te dire, ce que je peux pas écrire. Faudrait que j'invente des mots qui existent pas dans le dico... Vachement beaucoup!





## Valorisations scientifiques

### Articles publiés

Jamoneau, A., Passy, S.I., Soininen J., **Leboucher, T.**, & Rosebery, J. (2018). Beta-diversity of diatom species and ecological guilds : response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology*, **63** (1), 62-73.

Soininen, J., Jamoneau, A., Rosebery, J., **Leboucher, T.**, Wang, J., Kokociński, M., & Passy, S.I. (2018). Stream diatoms exhibit weak niche conservation along global environmental and climatic gradients. *Ecography*.

Passy, S.I., Larson, C.A., Jamoneau, A., Budnick, W.R., Heino, J., **Leboucher, T.**, Tison-Rosebery, J., & Soininen, J. (2018). The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. *The American naturalist*, **192** (5).

### Articles soumis ou en préparation

**Leboucher, T.**, Budnick, W.R., Passy, S.I., Bourtry, S., Jamoneau, A., Soininen, J., Vyverman, W., & Tison-Rosebery, J. (*Accepted*). Diatom  $\beta$ -diversity in streams increases with spatial scale and decreases with nutrient enrichment across regional to sub-continental scales. *Journal of Biogeography*.

Budnick, W.R., **Leboucher, T.**, Pound, K., Jamoneau, A., Tison-Rosebery, J., Soininen, J., Pajunen, V., Belliard, J., Campeau, S., Tales, E., & Passy, S.I. (*under review*). Local and regional drivers of taxonomic homogenization in stream communities along a land use gradient. *Global Ecology and Biogeography*.

**Leboucher, T.**, Tison-Rosebery, J., Budnick, W.R., Jamoneau, A., Vyverman, W., Soininen, J., Boutry, S., & Passy, S.I. (*in prep. for Journal of Applied Ecology*). Asymmetric eigenvector maps and co-occurrence patterns provide robust evidence for mass effect influence on stream diatoms.

**Leboucher, T.**, Mignien, L., Wach, M., Boutry, S., Jamoneau, A., & Tison-Rosebery, J. (*in prep. for Ecological Indicators*). Taking into account mass effects processes in bioindication allows a more accurate bioassessment of water quality.

## Communications orales

Jamoneau, A., **Leboucher, T.**, Bottin, M., Soininen, J., Passy, S.I., & Tison-Rosebery, J. Benthic diatom metacommunities : an illustrated overview of processes and patterns. *4<sup>th</sup> Biofilm Workshop*, 7-8 April 2016, Girona, Spain. Lecture.

**Leboucher, T.**, Budnick, W.R., Passy, S.I., Bourtry, S., Jamoneau, A., Soininen, J., Vyverman, W., & Tison-Rosebery, J. Shifts in beta-diversity determinants from local to global scales : the case study of aquatic diatom communities. *10<sup>th</sup> Symposium for European Freshwater Sciences*, 2-7 July 2017, Olomouc, Czech Republic. Oral communication.

Jamoneau, A., Passy, S.I., Soininen, J., **Leboucher, T.**, & Tison-Rosebery, J. From up- to downstream : beta-diversity patterns of diatoms communities in freshwater ecosystems. *10<sup>th</sup> Symposium for European Freshwater Sciences*, 2-7 July 2017, Olomouc, Czech Republic. Oral communication.

**Leboucher, T.**, Budnick, W.R., Passy, S.I., Bourtry, S., Jamoneau, A., Soininen, J., Vyverman, W., & Tison-Rosebery, J. Drivers of stream diatom beta-diversity from local to global scales. *Functional Ecology and Environment conference*, 11-12 July 2017, Toulouse, France. Oral communication.

**Leboucher, T.**, Budnick, W.R., Passy, S.I., Bourtry, S., Jamoneau, A., Soininen, J., Vyverman, W., & Tison-Rosebery, J. Drivers of stream diatom beta-diversity from local to global scales. *1<sup>st</sup> International Conference on Community Ecology*, 28-29 September 2017, Budapest, Hungary. Oral communication.

Soininen, J., Jamoneau, A., Rosebery, J., **Leboucher, T.**, Wang, J., Kokociński, M., & Passy, S.I. Niche conservatism among stream diatoms. *International Conference for Ecological Sciences*, 22-25 October 2018, Rennes, France. Oral communication.





# Table des matières

|                                                                                                                                                                           |           |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Introduction</b>                                                                                                                                                       | <b>19</b> |
| <b>1 Écologie des communautés de diatomées : état des connaissances</b>                                                                                                   | <b>25</b> |
| 1.1 Les diatomées . . . . .                                                                                                                                               | 25        |
| 1.1.1 Présentation générale . . . . .                                                                                                                                     | 25        |
| 1.1.2 Utilisation en bio-indication . . . . .                                                                                                                             | 28        |
| 1.2 Écologie des communautés . . . . .                                                                                                                                    | 31        |
| 1.2.1 Processus déterministes . . . . .                                                                                                                                   | 31        |
| 1.2.2 Les processus stochastiques . . . . .                                                                                                                               | 32        |
| 1.2.3 Les quatre paradigmes du concept de méta-communautés . . . . .                                                                                                      | 33        |
| 1.2.4 Quelles sont les connaissances actuelles sur l'influence des différents processus agissant sur la structuration des communautés de diatomées benthiques ? . . . . . | 36        |
| <b>2 Méthodologie : description des bases de données et des principales méthodes utilisées</b>                                                                            | <b>41</b> |
| 2.1 Bases de données . . . . .                                                                                                                                            | 41        |
| 2.1.1 Base de données Pandore . . . . .                                                                                                                                   | 41        |
| 2.1.2 Base de données du <i>National Water-Quality Assessment Program (NAWQA)</i> . . . . .                                                                               | 43        |
| 2.1.3 Paramètres climatiques . . . . .                                                                                                                                    | 43        |
| 2.1.4 Autres paramètres . . . . .                                                                                                                                         | 44        |
| 2.2 Principales méthodes utilisées . . . . .                                                                                                                              | 44        |
| 2.2.1 Mesures de la biodiversité . . . . .                                                                                                                                | 44        |
| 2.2.2 Cadre multi-échelles explicite : développement de la méthode des fenêtres paysagères . . . . .                                                                      | 47        |
| 2.2.3 Étude de l'influence relative des processus déterministes et stochastiques . . . . .                                                                                | 48        |
| 2.2.4 Influence des processus de dispersion et modélisation de la distribution des espèces . . . . .                                                                      | 50        |

|          |                                                                                                                                                                                           |            |
|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>3</b> | <b>Influences des processus déterministes et stochastiques sur la structuration des communautés de diatomées benthiques : approche multi-échelle et impact des perturbations humaines</b> | <b>55</b>  |
| 3.1      | Contexte et objectifs de l'étude . . . . .                                                                                                                                                | 55         |
| 3.2      | Principaux résultats . . . . .                                                                                                                                                            | 56         |
|          | Diatom $\beta$ -diversity in streams increases with spatial scale and decreases with nutrient enrichment across regional to sub-continental scales . . . . .                              | 57         |
| <b>4</b> | <b>Développement méthodologique pour la mise en évidence de l'influence de l'effet de masse sur la distribution des espèces</b>                                                           | <b>103</b> |
| 4.1      | Contexte et objectifs de l'étude . . . . .                                                                                                                                                | 103        |
| 4.2      | Principaux résultats . . . . .                                                                                                                                                            | 104        |
|          | Asymmetric eigenvector maps and co-occurrence patterns provide robust evidence for mass effect influence on stream diatoms . . . . .                                                      | 106        |
| <b>5</b> | <b>Influence de l'effet de masse en bio-indication : exemple de l'Indice Biologique Diatomée</b>                                                                                          | <b>141</b> |
| 5.1      | Contexte et objectifs de l'étude . . . . .                                                                                                                                                | 141        |
| 5.2      | Principaux résultats . . . . .                                                                                                                                                            | 142        |
|          | Taking into account mass effect processes in bioindication allows a more accurate bioassessment of water quality . . . . .                                                                | 143        |
| <b>6</b> | <b>Apports, limites et perspectives du travail réalisé</b>                                                                                                                                | <b>161</b> |
| 6.1      | Bilan général . . . . .                                                                                                                                                                   | 161        |
| 6.1.1    | Influences des processus déterministes et stochastiques sur la structuration des communautés de diatomées benthiques . . . . .                                                            | 161        |
| 6.1.2    | Développement méthodologique pour la mise en évidence de l'influence de l'effet de masse sur la distribution des espèces . . . . .                                                        | 163        |
| 6.1.3    | Influence de l'effet de masse en bio-indication : exemple de l'Indice Biologique Diatomée . . . . .                                                                                       | 164        |
| 6.2      | Perspectives de recherche . . . . .                                                                                                                                                       | 165        |
| 6.2.1    | Base de données . . . . .                                                                                                                                                                 | 165        |
| 6.2.2    | Occupation des sols et modèles hydrologiques . . . . .                                                                                                                                    | 165        |
| 6.2.3    | Structuration des communautés . . . . .                                                                                                                                                   | 167        |
|          | <b>Références</b>                                                                                                                                                                         | <b>171</b> |





# Introduction

Les paysages continentaux subissent depuis plusieurs siècles une transformation profonde, passant de grandes forêts continues et d'étendues inexploitées, à des zones agricoles et à des zones urbaines en constante expansion (DeFries et al., 2004; McKinney, 2006; Newbold et al., 2015). Ces transformations entraînent un déclin de la biodiversité au niveau mondial, qui ne fait que s'accélérer depuis plusieurs décennies (McGill et al., 2015; Newbold et al., 2015; Vitousek et al., 1997), se traduisant aujourd'hui par une homogénéisation des communautés dans les milieux modifiés par l'Homme (McKinney and Lockwood, 1999; Solar et al., 2015; Vallejos et al., 2016). Cet accroissement de la similarité entre les communautés ne touche pas seulement quelques espèces, mais est bel et bien un phénomène généralisé documenté chez de nombreux taxa (Blair, 2001; Gossner et al., 2016; Karp et al., 2012; Sotherton, 1998). Appréhender les mécanismes sous-jacents à ces changements est certainement l'une des facettes les plus intéressantes de l'écologie des communautés. Une communauté observée à un instant donné est ainsi le résultat de la combinaison d'une multitude de facteurs, biotiques et abiotiques, et démêler l'influence relative de chacun de ces facteurs représente, encore aujourd'hui, un défi majeur pour bien comprendre la dynamique d'assemblage des espèces et en conséquence mieux assurer le suivi de la biodiversité et sa conservation. Ces questionnements s'inscrivent pleinement dans le domaine de l'écologie méta-communautaire qui considère les communautés locales en tant qu'assemblages d'espèces liés les uns aux autres par la dispersion (Leibold et al., 2004).

La variation locale de la composition en espèces (diversité  $\alpha$ ), au sein d'une multitude de communautés composées à partir d'un *pool* régional d'espèces (diversité  $\gamma$ ), a été surnommée diversité  $\beta$  par Robert J. Whittaker il y a plus de cinquante ans (1960; 1972). La diversité  $\beta$  a été définie à l'origine comme l'ampleur du changement dans la composition entre deux communautés. Ce concept s'est depuis beaucoup transformé, évoluant d'une simple valeur scalaire ( $\beta = \gamma/\bar{\alpha}$ ) en de nombreuses définitions mathématiques offrant de nombreux et puissants outils pour étudier les différents aspects de ce concept de variabilité spécifique (Socolar et al., 2016; Tuomisto, 2010a,b). Le déclin rapide de la biodiversité au cours de l'Anthropocène a conduit les écologues à s'interroger sur les moteurs de la diversité  $\beta$  dynamisant ainsi son étude. Des processus déterministes, comme les filtres environnementaux et les relations biotiques (classiquement résumés sous le terme de *species sorting*), et des facteurs stochastiques, tels que les processus de dispersion, ont ainsi été identifiés comme des mécanismes agissant sur les patrons observés de diversité  $\beta$  (Condit et al., 2002; Legendre et al., 2005). Ces études ont égale-

ment mis en évidence que l'influence relative de ces processus était contrainte et variait avec l'échelle spatiale considérée, et qu'il apparaissait de fait nécessaire de prendre en compte cette dimension pour mieux comprendre comment les différents processus locaux et régionaux affectaient la diversité  $\beta$  (Huston, 1994; Mac Nally et al., 2004; Scott et al., 2002).

La prise en compte de processus stochastiques, comme la dispersion, sur la structuration de communautés biotiques est encore relativement nouvelle en écologie des communautés et n'a véritablement pris de l'ampleur qu'à l'avènement de l'écologie méta-communautaire au début du vingt-et-unième siècle (Leibold et al., 2004). Parmi les processus de dispersion, l'effet de masse et la dispersion limitée sont les deux mécanismes les plus régulièrement évoqués pour expliquer la structuration des communautés (Bottin et al., 2016; Jamoneau et al., 2018; Soininen, 2007). L'effet de masse est défini comme la persistance d'une espèce dans un habitat qui lui est défavorable grâce à l'apport de nouveaux individus via l'immigration. La dispersion limitée, quant à elle, est définie par l'incapacité d'une espèce à occuper tous les habitats qui lui sont favorables en raison de capacités de dispersion limitées ou de barrières physiques, naturelles ou non (chaînes de montagnes ou barrages par exemple), empêchant sa dispersion.

Ces dernières années, de nombreuses études se sont concentrées sur la délicate question de démêler l'influence relative de ces deux mécanismes sur la structuration des méta-communautés. Si la dispersion limitée est relativement bien appréhendée, par le biais d'approches expérimentales (Ehrlén and Eriksson, 2000) ou de modélisation (Svenning et al., 2008), l'effet de masse reste encore aujourd'hui beaucoup plus théorique. Pour les organismes à dispersion active, tels certains macro-invertébrés, l'utilisation de traits spécifiques signant les capacités de dispersion des espèces peuvent être utilisés pour attester de la présence de ce mécanisme (Brown and Swan, 2010). Dans le cas des organismes à dispersion passive, ce n'est malheureusement pas possible et seules des preuves indirectes sont utilisées pour évoquer le rôle de l'effet de masse sur la structuration des communautés. Les écologues restent ici confrontés à un véritable verrou scientifique : comment déterminer la niche d'une espèce potentiellement présente dans des conditions qui lui sont défavorables ?

La bonne compréhension de ce mécanisme est pourtant importante dans le cadre du déclin mondial de la biodiversité. L'effet de masse a en effet un fort pouvoir d'homogénéisation des communautés puisqu'il permet à une espèce de sortir de sa niche et d'être présente largement dans l'écosystème. Ce phénomène peut également impacter les méthodes de bio-indication de la qualité des milieux (Jamoneau et al., 2018). La présence d'espèces dans des habitats *a priori* défavorables, à cause de l'effet de masse, masque potentiellement l'effet des conditions locales induisant ainsi une mauvaise évaluation du milieu (Heino et al., 2015; Leibold et al., 2004). Il semble évident qu'une meilleure intégration de ce mécanisme est aujourd'hui nécessaire, afin d'améliorer le développement d'indices d'évaluation plus efficaces, intégrant à la fois les processus environnementaux et de dispersion.

C'est dans ce contexte scientifique que s'inscrit notre travail de recherche dont les principaux objectifs peuvent être énumérés comme suit :

- déterminer l'influence relative des processus déterministes et stochastiques dans la structuration des méta-communautés, notamment dans le contexte actuel où l'Homme a un fort impact sur les écosystèmes, et ce, à différentes échelles spatiales ;
- tenter d'apporter une méthode capable de mettre en évidence et quantifier l'influence des processus de dispersion dans la structuration des communautés ;
- enfin, étudier l'impact de la prise en compte de ces processus sur les méthodes de bio-indication de la qualité des milieux.

Les communautés de diatomées benthiques de rivières semblent être un modèle biologique particulièrement adapté pour explorer ces objectifs de recherche. Ces algues unicellulaires jouent un rôle important dans les écosystèmes aquatiques, notamment en termes de production primaire et dans un certain nombre de cycles biogéochimiques. Au même titre que les plantes supérieures, il est aujourd'hui reconnu que les diatomées sont influencées par les filtres environnementaux (Finlay, 2002; Verleyen et al., 2009). Des études récentes tendent également à montrer qu'elles sont influencées par des processus de dispersion (Göthe et al., 2013; Jamoneau et al., 2018; Soininen et al., 2016; Vanormelingen et al., 2008; Vyverman et al., 2007), mais cela n'a jamais été étudié dans un cadre spatial étendu, de l'échelle locale à l'échelle intercontinentale, comme nous le proposons ici. Enfin, ce sont des organismes à dispersion passive, largement utilisés en bio-indication de la qualité de l'eau. Ces organismes offrent donc un cadre idéal pour notre travail.

Pour le mener à bien, nous disposons d'une large base de données françaises et américaines contenant plus de 20 000 relevés diatomiques réalisés entre le début des années 90 et aujourd'hui. Ces données sont la compilation des données récoltées annuellement par les Agences de l'Eau en France, et par le programme d'évaluation nationale de la qualité de l'eau aux États-Unis (*National Water-Quality Assessment Program*), et représentent une quantité de données inédite augmentant la fiabilité de nos résultats et l'intérêt de ce travail de recherche.

Ce mémoire est construit en six chapitres :

- le **premier chapitre** introduit des éléments généraux concernant les communautés de diatomées benthiques, ainsi que les méthodes de bio-indication qui les utilisent. Il expose ensuite l'état de nos connaissances concernant les mécanismes structurant les méta-communautés de diatomées, ce qui nécessite d'introduire en amont des notions fondamentales d'écologie des communautés.

- le **second chapitre** introduit les principales méthodes utilisées au cours de cette thèse.
- le **troisième chapitre** expose nos résultats concernant l'influence relative des processus déterministes et stochastiques sur la structuration des méta-communautés diatomiques, à différentes échelles spatiales et en prenant en compte l'impact de perturbations anthropiques (eutrophisation).
- le **quatrième chapitre** s'attache à mettre en évidence le rôle précis de l'effet de masse dans la structuration des communautés diatomiques.
- le **cinquième chapitre** replace nos résultats des second et troisième chapitres dans le cadre de la bio-indication de la qualité de l'eau, et s'attache à mesurer l'impact de la prise en compte de l'effet de masse sur la bio-indication.
- enfin, le **sixième chapitre** discute des principaux résultats obtenus et ouvre les perspectives de recherche qui potentiellement en découlent.





# 1 Écologie des communautés de diatomées : état des connaissances

Ce chapitre présente des informations générales concernant les communautés diatomiques et leur utilisation en bio-indication, ainsi que l'état des connaissances concernant l'écologie des communautés, et en particulier les problématiques actuelles liées aux micro-organismes.

## 1.1 Les diatomées

### 1.1.1 Présentation générale

Les diatomées sont des algues brunes (Chromophytes), microscopiques et unicellulaires appartenant à la classe des Bacillariophycées. Leur taille varie de moins de 10  $\mu\text{m}$  à plus de 500  $\mu\text{m}$ . Elles sont composées d'organites différenciés au sein de leur cellule (noyau, mitochondries, chloroplastes, etc.) et sont structurées au sein d'un exosquelette siliceux. Cette structure, appelée frustule, est composée de deux valves emboîtées (épi- et hypo-valves) unies par une ceinture connective, et présente des formes et ornements extrêmement variés à la base des classifications taxonomiques.

Apparues au cours du Jurassique, il y a 200 millions d'années environ, la grande majorité des diatomées sont autotrophes et possèdent de faibles besoins en lumière et en humidité, ce qui explique leur développement dans des milieux aussi variés que l'eau (douce, salée et saumâtre, courante, stagnante, suintante et intermittente), l'air (aérosols) et le sol (sols humides, cavernes, etc.). Les estimations actuelles font état de plus de 30 000 espèces découvertes et de plus de 100 000 potentielles (Mann and Vanormelingen, 2013).

## Éléments de biologie

Les diatomées sécrètent un mucilage à la surface de leur frustule. Ce mucilage joue un rôle majeur dans la mobilité des diatomées, leur adhésion au substrat, dans leur capacité à former des colonies et contre la dessiccation (Hoagland et al., 1993).

La multiplication des diatomées est principalement végétative. Chaque cellule mère donne deux cellules filles par bipartition suite à une mitose classique (Fig. 1.1). Chaque cellule fille conserve l'une des deux valves de la cellule mère qui devient l'épivalve, et fabrique une hypovalve de plus petite taille venant s'emboîter dans la première. La taille des cellules filles est donc asymétrique, l'une conservant la taille de la cellule mère et l'autre devenant plus petite. Lorsque la taille des diatomées atteint une limite critique, l'auxospore intervient. C'est un phénomène de reproduction sexuée caractéristique des Bacillariophycées pendant lequel chaque cellule mère subit une méiose produisant quatre noyaux haploïdes dont seul un va subsister, fonctionner comme un gamète, et croître jusqu'à sa taille maximale, permettant à l'individu de retrouver la taille initiale de son espèce.

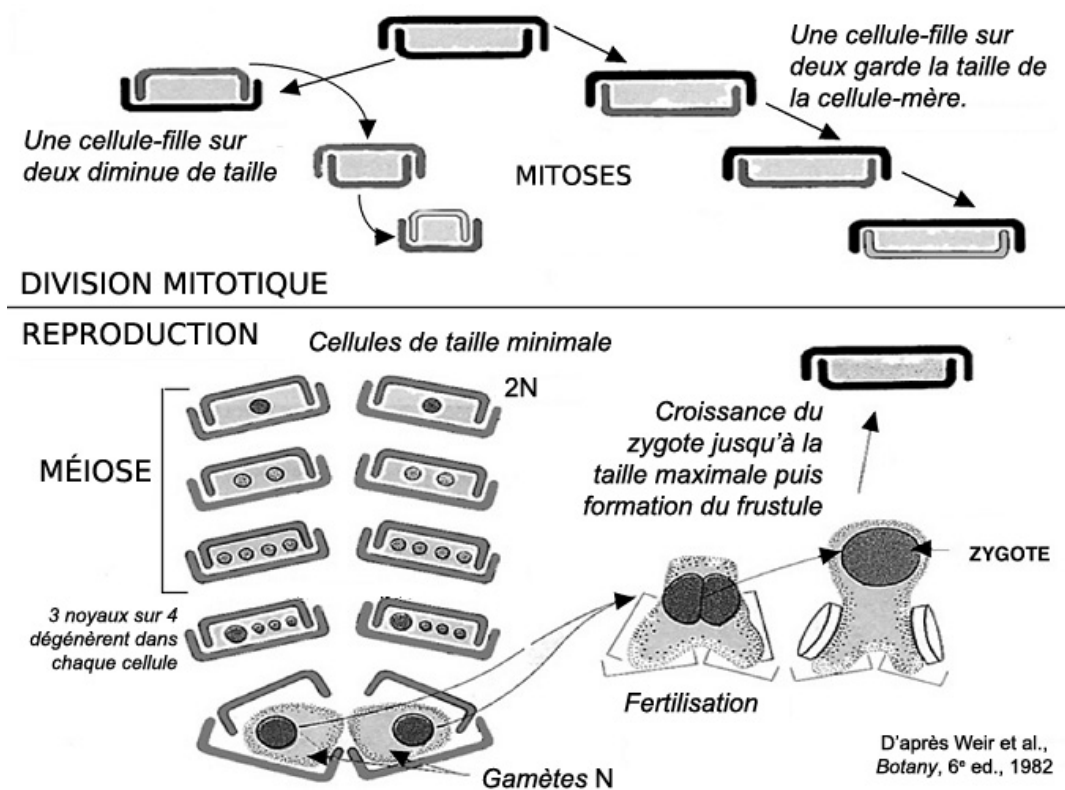


FIGURE 1.1 – Représentation schématique des cycles de multiplication végétative et de la reproduction méiotique chez les diatomées.

## Classification

La systématique des diatomées se base sur l'observation des caractères structuraux et morphologiques des frustules. Ces observations sont tributaires de la performance technique des outils de microscopie, ce qui explique la grande complexité de la systématique des diatomées et sa constante évolution. Actuellement, tous les auteurs sont d'accord pour subdiviser la classe des Bacillariophycées en deux sous-classes :

- Les Centrophycidées, ou diatomées centriques. De forme ronde, elles sont souvent comparées à des boîtes de Pétri.
- Les Pennatophycidées, ou diatomées pennées. De forme allongée, elles portent le nom de pennées en rapport aux stries disposées perpendiculairement à l'axe longitudinal rappelant la forme des plumes d'oiseaux.

Cette classification basée sur des critères morphologiques, fait relativement rare chez les micro-organismes, est aujourd'hui complétée par l'avancée de méthodes moléculaires telles que le métabarcoding. La complexité de l'approche taxonomique présente en effet quelques inconvénients, notamment le besoin de personnel qualifié. Cependant, l'application de ces méthodes moléculaires ne semble pas si simple à mettre en œuvre et démontre à ce jour un manque de concordance avec la taxonomie "classique". De plus, la très grande diversité des diatomées ne permet pas pour l'instant une analyse exhaustive des différents marqueurs moléculaires pour un nombre acceptable d'espèces.

## Écologie générale

Il existe deux grandes formes de vie chez les diatomées : les diatomées planctoniques et les diatomées périphtiques :

- Les diatomées planctoniques vivent libres dans l'eau et ne sont pas liées au substrat. Leur forte mobilité leur permet de coloniser l'ensemble des sections du réseau hydrographique.
- Les diatomées périphtiques, ou fixées, dont la classification repose sur les types de micro-habitats occupés par chaque espèce. Ainsi, l'épiphyton rassemble les espèces vivant à la surface des végétaux ; l'épipélon et l'endopélon désignent respectivement les espèces vivant à la surface et dans les sédiments ; l'endopsammon fait référence aux espèces vivant à la surface des grains de sable, et l'épilithon rassemble les espèces vivant sur les substrats durs tels que pierres, galets, blocs qu'ils soient d'origine naturelle ou non. C'est ce dernier compartiment qui a été

principalement pris en compte lors de ce travail. Enfin, le terme périphyton désigne plus largement toutes les espèces vivant fixées sur des objets immergés. Ce sont ces espèces (plus particulièrement celles de l'épilithon) qui sont utilisées en bio-indication.

La structure formée par les diatomées du périphyton, liées les unes aux autres par la sécrétion de mucilage notamment, est appelée biofilm. Celui-ci peut également comprendre d'autres types d'organismes tels que des algues filamenteuses, des bactéries, etc. La grande diversité des formes et des tailles de diatomées composant le périphyton permet d'expliquer la structure irrégulière et complexe du biofilm. Le biofilm peut ainsi exposer des différences structurelles aussi marquées que celles séparant les pelouses des forêts, et des gradients verticaux de nutriments y sont par exemple observés (Passy, 2007). Devant cette complexité, Passy (2007) a défini trois guildes, aujourd'hui très utilisées pour décrire la structure tridimensionnelle du biofilm :

- *Low profile* : espèces adhérentes au substrat vivant dans les couches inférieures du biofilm. Elles sont souvent considérées comme des espèces pionnières, base pour l'ensemble du biofilm.
- *High profile* : espèces érigées formant des colonies leur permettant d'atteindre les couches supérieures du biofilm.
- *Motile* : espèces pouvant se déplacer au sein du biofilm.

### 1.1.2 Utilisation en bio-indication

L'intérêt d'utiliser des organismes vivants pour évaluer la qualité du milieu repose sur le fait qu'ils permettent, à l'inverse des mesures chimiques et physiques ponctuelles qui ne sont qu'une photographie à un instant  $t$ , d'intégrer globalement la qualité de l'eau sur de relativement longues périodes de temps.

Les diatomées présentent une grande sensibilité aux conditions du milieu, variant en fonction des espèces. Elles semblent particulièrement sensibles aux gradients liés à la balance ionique des eaux (Potapova and Charles, 2002; Van Dam et al., 1994) et aux concentrations de nutriments (en particulier aux phosphates et nitrates, Soininen 2007; Soininen et al. 2016). Les diatomées sont également sensibles aux gradients altitudinaux (Soininen et al., 2016), en lien avec les conditions hydrodynamiques, d'oxygène dissous et de températures.

Cette sensibilité aux conditions du milieu, leurs fortes abondances qui permettent un échantillonnage fiable et en quantité suffisante, et leur grande diversité d'espèces font des diatomées d'excellents indicateurs de la qualité de l'eau, notamment en rivière, compartiment physique étudié lors de cette thèse.

## Présentation de l'Indice Biologique Diatomée (IBD)

Les méthodes de bio-indication basées sur les diatomées sont très nombreuses et plus d'une vingtaine ont par exemple été répertoriées en Europe (Prygiel et al., 1999). La méthode normalisée en France en rivière est l'IBD (NFT 90-354, AFNOR 2007). Il repose sur la méthode des moyennes pondérées qui implique une description préalable des préférences des espèces en termes de conditions environnementales. Dans le cadre de l'IBD, ces valeurs ont été obtenues après l'analyse numérique de bases de données (Coste et al., 2009; Prygiel and Coste, 2000) et sont exprimées en probabilité de présence de l'espèce en fonction de sept classes prédéfinies de qualité de l'eau. Le calcul se base sur la formule de Zelinka and Marvan (1961) et peut se décliner comme suit (équations 1.1 et 1.2) :

$$IBD = \sum_{j=1}^7 F_j \cdot j \quad (1.1)$$

Où

$$F_j = \frac{\sum_{i=1}^n A_i V_i P_{ij}}{\sum_{i=1}^n A_i V_i} \quad (1.2)$$

Et où  $A_i$  est l'abondance de l'espèce  $i$ ,  $V_i$  est son degré de sténocécie et  $P_{ij}$  est l'affinité de l'espèce  $i$  avec la classe  $j$ .

La définition de ces probabilités de présence provient de descriptions de profils écologiques des taxons (Coste et al., 2009), qui peuvent être illustrés comme sur la figure 1.2.

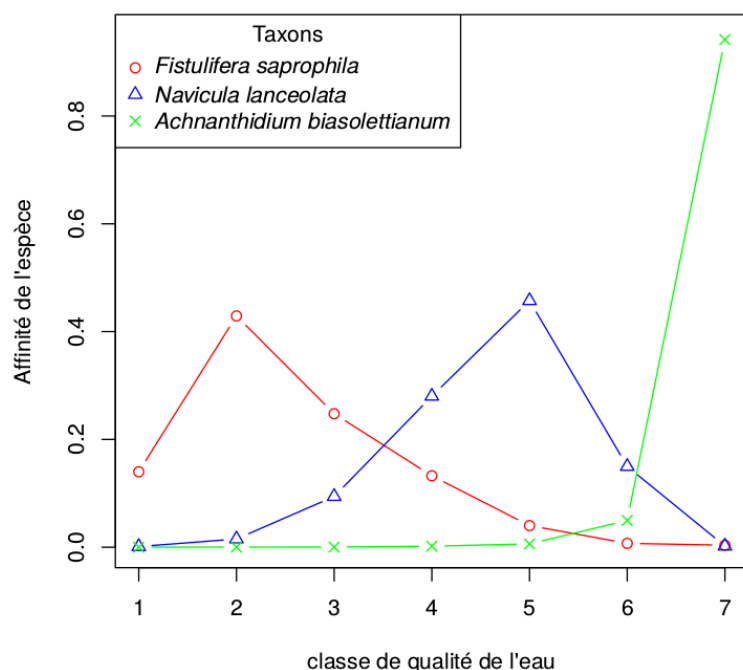


FIGURE 1.2 – Trois exemple de profils écologiques définis pour le calcul de l'IBD et présentés dans la thèse de M. Bottin (Bottin, 2012). *Fistulifera saprophila* est une espèce sensible (forte sténocéie) aux conditions du milieu capable de se développer dans des eaux de mauvaise qualité. *Navicula lanceolata* est une espèce peu sensible aux conditions du milieu (euryèce) présente sur une grande partie du gradient de qualité de l'eau. *Achnanthidium biasolettianum* est une espèce très sensible aux conditions du milieu (forte sténocéie), uniquement présente sur les classes 6 et 7 correspondant aux eaux de très bonne qualité.

La probabilité  $P_{ij}$  est calculée selon l'équation 1.3 :

$$P_{ij} = \frac{N_i A_i}{S_i \cdot \sum_{j=1}^7 \frac{N_i A_i}{S_j}} \quad (1.3)$$

Où  $N_i$  est l'occurrence de l'espèce  $i$ ,  $A_i$  est l'abondance de l'espèce  $i$ ,  $S_j$  est le nombre de sites au sein de la classe de qualité  $j$ .

Les notes obtenues sont ensuite transformées en classe d'état (cinq au total) allant du très mauvais état au très bon état, en accord avec la Directive Cadre européenne sur l'Eau (DCE). Afin de tenir compte des spécificités régionales découvertes par les travaux de Tison et al. (2005), le calcul de classe d'état n'est pas le même en fonction des régions considérées (JORF, 2015).

## 1.2 Écologie des communautés

L'écologie des communautés est l'une des disciplines de l'écologie (étant elle-même une branche de la biologie) qui s'intéresse aux interactions entre les organismes vivant dans un habitat donné, et leur environnement, ainsi qu'aux interactions des organismes entre eux. Cette définition reste toutefois très théorique puisqu'il est pour ainsi dire impossible de considérer en même temps l'ensemble des organismes vivants, ne serait-ce que sur un habitat très restreint. Cette définition est donc classiquement réduite à un groupe d'organismes, ici les diatomées. Nous passerons donc en revue les principaux concepts et mécanismes agissant sur les communautés biotiques, en faisant le point sur l'état des connaissances concernant les diatomées, organismes clefs utilisés au cours de cette thèse.

### 1.2.1 Processus déterministes

#### **Théorie des niches et filtres environnementaux**

Le concept de niche a été défini pour la première fois par Grinnell en 1917 comme l'ensemble des traits qui permettent à une espèce de répondre à ses besoins environnementaux, d'échapper à ses ennemis naturels, ainsi que tout autre facteur influençant ses taux de natalité et de mortalité. Hutchinson (1957) redéfinira plus tard ce concept pour lui donner son sens actuel d'hypervolume dans un espace à  $n$  dimensions,  $n$  étant le nombre de facteurs pouvant influencer la distribution de l'espèce.

Bien que ce concept soit bien souvent salué comme l'un des concepts les plus importants en écologie de communautés, par exemple dans Chase and Leibold (2003), les études portant sur les niches écologiques sont considérées tautologiques par une partie de la communauté puisqu'il ne semble toujours pas possible de définir clairement le nombre  $n$  de dimensions de la niche (Clark et al., 2010, 2007). De plus, Siepielski and McPeck (2010) ont récemment mis en évidence que les preuves attestant que la coexistence stable des espèces peut être expliquée par des différences de niches, sont en réalité très rares.

Si la niche seule ne peut pas expliquer la cohabitation entre les espèces, les  $n$  dimensions de l'hypervolume d'Hutchinson peuvent être considérées comme autant de filtres environnementaux agissant sur la structuration de l'espèce considérée. Une espèce pour pouvoir s'installer et perdurer dans un habitat doit être capable d'en supporter toutes les contraintes. Là où les gradients environnementaux persistent à des échelles temporelles suffisamment grandes, les espèces développent ainsi des caractéristiques particulières adaptées à l'environnement local, appelées traits (Gutiérrez-Cánovas et al., 2013). Ces

traits de vie sont le produit de la sélection naturelle et sont classiquement séparés en deux parties. D'un côté les traits biologiques correspondent aux stratégies de reproduction, à la croissance des individus, etc., et de l'autre les traits écologiques correspondent aux préférences environnementales de l'espèce en termes ressources, de température, de pH, etc. Ces traits écologiques ne sont en réalité que la traduction de traits biologiques.

## **Les interactions biotiques**

L'unique considération des conditions environnementales ne peut pas toujours permettre d'expliquer la présence ou non d'une espèce dans un milieu qui lui est pourtant favorable. En effet, les relations biotiques entre les espèces, qu'elles soient positives (facilitation, complémentarité de niche) ou négatives (prédation, compétition) déterminent la niche réalisée des espèces, en opposition à la niche fondamentale définie par les seuls filtres environnementaux. La compétition apparaît comme un processus important puisqu'elle permettrait d'expliquer la relative pauvreté de la diversité à l'échelle locale par l'exclusion mutuelle des espèces (MacArthur and Levins, 1967).

### **1.2.2 Les processus stochastiques**

#### **La théorie neutraliste**

A l'inverse de la théorie des niches, la théorie neutraliste introduite par Hubbell (2001) assume que des processus stochastiques seuls sont capables d'expliquer la distribution des espèces. Hubbell y réfute donc l'importance du concept de niche considérant toutes les espèces comme étant équivalentes en termes de traits biologiques, et explique la structuration des communautés grâce aux taux de dispersion, de spéciation et d'extinction.

L'introduction de cette nouvelle théorie a suscité un débat important parmi les écologues et a bien failli rendre obsolète le concept de niche (Letten et al., 2017). De ce débat sont ressortis de nouveaux concepts considérant que ces deux théories n'étaient pas forcément antinomiques, comme le concept de méta-communautés (Leibold et al., 2004).

## Les processus de dispersion

Deux principaux processus spatiaux agissent sur la structuration des communautés (Heino et al., 2015; Mouquet and Loreau, 2003; Shmida and Wilson, 1985) :

**La dispersion limitée** peut être définie par l'incapacité d'une espèce à occuper tous les habitats qui lui sont favorables en raison de capacités de dispersion limitées, ou de barrières physiques (chaînes de montagnes ou barrages par exemple) empêchant sa dispersion. Ce processus est habituellement décrit à de larges échelles spatiales (supérieures au bassin versant) et a un impact positif sur la diversité, puisqu'il entraîne une différenciation artificiellement plus importante entre les sites partageant des conditions environnementales communes. C'est d'ailleurs cet effet sur la diversité qui permet de le détecter puisqu'il est souvent mis en évidence par le biais de signaux spatiaux à large échelle couplés à une forte diversité  $\beta$ . Ces résultats pourraient cependant être expliqués par d'autres processus, comme la compétition et/ou la présence de gradients environnementaux spatialisés non mesurés, ce qui peut rendre sa détection compliquée.

**L'effet de masse** peut quant à lui être défini comme la persistance d'une espèce dans un habitat qui lui est défavorable grâce à l'apport de nouveaux individus via l'immigration, et tend à homogénéiser des communautés puisque des sites avec des environnements locaux différents pourront ainsi partager les mêmes espèces. Ce phénomène est plutôt décrit à de petites échelles spatiales et est tout aussi compliqué à détecter. En effet, sa présence est généralement attestée par le couplage de signaux spatiaux avec de faibles valeurs de diversité  $\beta$ . Comme dans le cas de la dispersion limitée, d'autres processus (par exemple la présence de gradients environnementaux spatialisés non mesurés ou des processus de complémentarité des niches) peuvent conduire à ces mêmes observations rendant la détection de l'effet de masse difficile.

Dans un cas comme dans l'autre, ces processus restent des modèles théoriques manquant pour le moment de méthodes pour attester de leur influence sur les communautés de façon convaincante.

### 1.2.3 Les quatre paradigmes du concept de méta-communautés

Le concept de méta-communautés a été décrit par Leibold et al. (2004) et désigne un ensemble de communautés reliées entre elles par des phénomènes de dispersion des espèces. L'article fondateur de ce concept (Leibold et al., 2004) y introduit quatre grands

paradigmes permettant de décrire les relations entre les communautés (schématiquement représentés sur la figure 1.3), en faisant appel aux processus précédemment décrits :

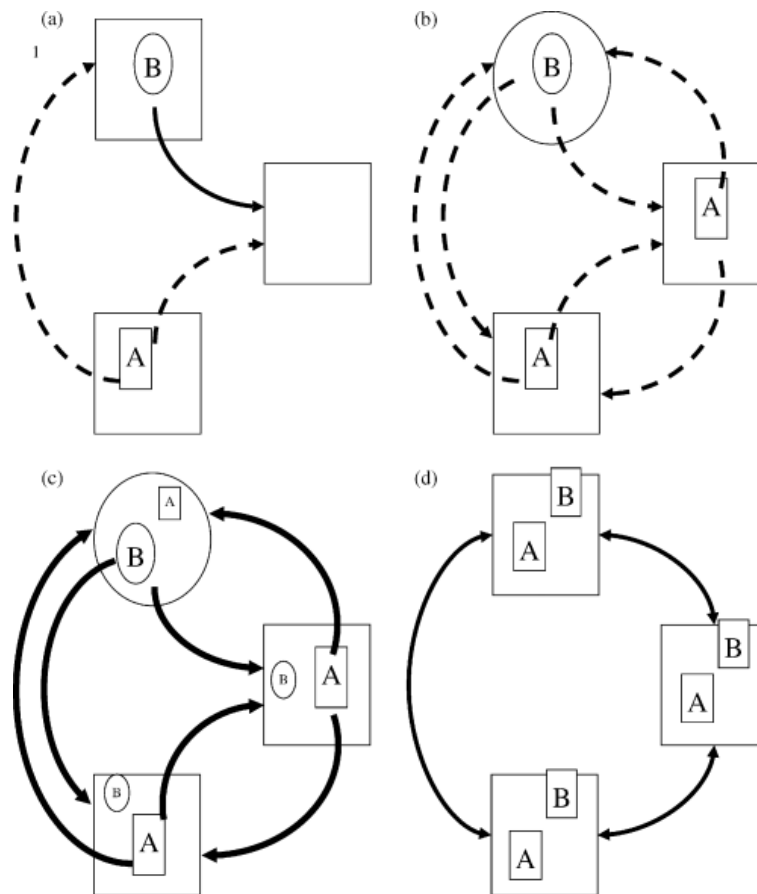


FIGURE 1.3 – Représentation schématique présentée par Leibold et al. (2004) des quatre paradigmes de la théorie des méta-communautés, (a) la dynamique de patch, (b) le *species sorting*, (c) l'effet de masse et (d) la dynamique neutre, pour deux espèces en compétition (populations A et B). Les flèches relient les populations sources aux sites de colonisation potentiels, représentés sous la forme de boîtes (ovales ou carrées). Les flèches pleines indiquent une dispersion plus élevée que les flèches en pointillé et si le mouvement est unidirectionnel (une seule tête) ou bidirectionnel (flèches à deux têtes). La compétition entre les deux espèces est indiquée par la taille des boîtes, une petite boîte indiquant que l'espèce est dominée.

**La dynamique de patch :** cette première approche assume que les communautés vivent sur des aires géographiquement délimitées appelées "patches", et subissent à la fois des processus déterministes et stochastiques. Cette approche emprunte à la théorie neutre l'importance des processus de dispersion et d'extinction, mais

n'exclue pas pour autant l'importance de la compétition entre les espèces. Les traits de vie des espèces sont alors un compromis entre capacités de dispersion et compétitivité. Sur la figure 1.3, ce paradigme (a) est présenté avec des conditions permettant la coexistence. Un compromis entre compétition et colonisation est illustré avec l'espèce A, bon compétiteur, et l'espèce B, bon colonisateur, rendant possible la colonisation du troisième patch par l'une ou l'autre de ces deux espèces.

**Le *species sorting*** : cette seconde approche s'appuie sur les théories d'adaptation des communautés le long de gradients environnementaux (voir Whittaker, 1962) et considère que les caractéristiques environnementales locales conditionnent les taux vitaux des populations et les interactions entre espèces (en lien avec la théorie des niches). Dans cette perspective, les différents patches sont considérés hétérogènes, et seules les conditions biotiques et abiotiques locales déterminent la composition spécifique. Sur la figure 1.3, ce paradigme (b) est représenté par deux espèces présentant deux niches distinctes (forme des boîtes) et où la dispersion n'est pas suffisante pour modifier leurs répartitions.

**L'effet de masse** : ce troisième paradigme intègre des éléments provenant des deux premiers. Chaque espèce se développe dans les sites où l'environnement leur est favorable et forme des populations sources pouvant disperser vers d'autres sites, dits puits. L'immigration de suffisamment d'individus depuis les populations sources permet le maintien de plus petites populations dans les sites puits même si les conditions environnementales n'y sont pas favorables pour l'espèce. Sur la figure 1.3, ce paradigme (c) entraîne la présence des deux espèces dans les habitats sources et puits. La forme des boîtes indique le type d'habitat, et la taille des lettres et des boîtes indique la taille des populations.

**La dynamique neutre** : les trois derniers paradigmes supposaient que les espèces diffèrent les unes des autres soit par leurs relations niches-environnement local, soit par leurs capacités de dispersion, ou encore soit par leur capacité à éviter les extinctions locales, entraînant de fait un jeu de compromis de traits ayant des conséquences aux échelles locale et régionale. En l'absence de telles différences entre les espèces, ce dernier paradigme, qui découle directement de la théorie d'Hubbell (2001), prédit des comportements méta-communautaires complètement différents où la dynamique de ces méta-communautés est influencée par des modèles aléatoires lents de changement de composition, dans l'espace et le temps. Sur la figure 1.3, ce paradigme (d) se traduit par la présence des deux espèces sur l'ensemble des sites. Les espèces seront amenées à disparaître progressivement de la région et à être remplacées par de nouvelles par le biais de processus de spéciation aléatoires.

#### 1.2.4 Quelles sont les connaissances actuelles sur l'influence des différents processus agissant sur la structuration des communautés de diatomées benthiques ?

Les communautés diatomiques n'ont longtemps été étudiées que sous le prisme de l'effet de l'environnement local ignorant totalement l'effet potentiel de processus de dispersion (Potapova and Charles, 2002; Tison et al., 2005). Celui-ci était considéré nul car la taille microscopique de ces organismes leur valait d'être considérés comme étant présents partout sans restrictions dues à des limites de dispersion. La structuration des communautés diatomiques apparaît aujourd'hui comme étant le résultat d'une interaction entre gradients biogéographiques, environnementaux et spatiaux générant une structuration à la fois déterministe et stochastique des *pool* d'espèces aux échelles locales et régionales (Jamoneau et al., 2018; Tonkin et al., 2015; Viana et al., 2016). Le *species sorting*, c'est-à-dire le couplage des filtres environnementaux et des interactions biotiques, ainsi que les processus de dispersion ressortent comme des facteurs particulièrement importants de la structuration des communautés (Chase and Leibold, 2003; Soininen, 2007), mais leur influence relative est peu connue et conditionnée par les échelles spatiales considérées. Cet effet des échelles spatiales a été au centre de l'intérêt des écologues et de nombreuses études ont ainsi montré que la distance moyenne entre les unités d'échantillonnage, leur configuration spatiale ou encore l'étendue spatiale de la région d'échantillonnage étaient des facteurs confondant limitant la comparaison des communautés entre les différentes zones géographiques étudiées (Conroy and Noon, 1996; Mac Nally et al., 2004; Noss, 1983; Wilson and Shmida, 1984). Des approches quantitatives multi-échelles sont donc recommandées pour mieux comprendre comment les processus locaux et régionaux affectent les communautés (Huston, 1994; Mac Nally et al., 2004), et pour identifier les échelles spatiales appropriées pour étudier les mécanismes qui structurent ces mêmes communautés (Mac Nally et al., 2004; Underwood and Chapman, 1996; Willis and Whittaker, 2002). Toutefois, peu d'études ont effectivement mis en œuvre des cadres spatiaux multi-échelles explicites pour explorer de telles relations, en particulier à larges échelles comme nous le proposons ici (échelle intercontinentale).

La littérature récente a donc mis en évidence que cet effet d'échelle était à relier à différents processus déterministes et stochastiques. Le *species sorting*, en particulier, façonne la composition des communautés le long des gradients environnementaux (Lack, 1976; Leibold et al., 2004) et donc le long de gradients spatiaux lorsque l'hétérogénéité environnementale augmente avec l'échelle. Là où les gradients environnementaux persistent à des échelles de temps géologiques, c'est-à-dire dans des conditions dites naturelles, l'adaptation locale des espèces à leur milieu (spéciation) entraîne une grande diversité de communautés à l'échelle régionale (Gutiérrez-Cánovas et al., 2013), et ainsi la diversité tend à augmenter avec l'échelle spatiale. Les processus spatiaux peuvent aussi influencer la diversité des communautés en fonction du taux de dispersion (Dong et al., 2016; Jamoneau et al., 2018; Soininen, 2007). L'effet de masse et la dispersion limitée sont

les deux principaux processus spatiaux agissant sur les communautés diatomiques, avec des effets antagonistes puisque le premier a tendance à homogénéiser les communautés alors que la dispersion limitée va au contraire avoir un impact positif sur la diversification des communautés. Ces deux processus sont cependant généralement détectés, tous les deux, par des signaux purement spatiaux et il est donc difficile de réellement attester de leurs influences relatives sur les communautés même si plusieurs études commencent à s'intéresser à cette question (voir Bottin et al., 2016; Goldenberg Vilar et al., 2014; Jamoneau et al., 2018). Ces deux processus restent donc encore mal compris et les preuves attestant de leurs influences encore trop indirectes.

La relation entre échelles spatiales et processus influant la diversité des communautés de diatomées est également influencée par les perturbations d'origine anthropique, qui causent une dégradation importante des écosystèmes aquatiques (Vörösmarty et al., 2010). Les perturbations anthropiques ont été associées à l'homogénéisation taxonomique mais les mécanismes sous-jacents à cette homogénéisation restent flous. Il semblerait que ce soit l'élimination des espèces sensibles (spécialistes, McKinney and Lockwood 1999; Olden and Poff 2003, 2004; Socolar et al. 2016) et le remplacement par quelques espèces généralistes (Gutiérrez-Cánovas et al., 2013), possédant une plasticité plus importante et des niches plus larges, qui soient à l'origine de ce phénomène. Dans le milieu naturel, la diversité d'espèces augmenterait, comme nous l'avons vu plus haut, avec l'échelle spatiale en raison d'une spécialisation accrue des habitats, favorisant ainsi une plus grande richesse et une prépondérance des spécialistes (Chase, 2014). Les perturbations anthropiques, en diminuant le nombre d'espèces spécialistes et en augmentant la prévalence des généralistes, pourraient alors causer une relation négative entre échelles spatiales et diversité des espèces.

En effet, le remplacement d'espèces spécialisées par des espèces généralistes est probablement le résultat d'interactions entre perturbations anthropiques et processus déterministes et stochastiques sous-jacents. De fait, l'une des conséquences de la disparition des espèces les plus sensibles aux conditions du milieu pourrait être la baisse de l'influence du *species sorting* sur les communautés, offrant une plus grande importance aux processus stochastiques, tels les processus de dispersion. L'importance des processus déterministes et stochastiques changeant avec l'échelle spatiale considérée, de la même manière que précédemment il semble difficile de prédire si les perturbations anthropiques sont capables de modifier, ou de conserver, l'influence relative des différents processus le long de gradients spatiaux.

Dans tous les cas, l'influence potentiellement croissante des processus stochastiques au détriment du *species sorting* à cause de la prévalence d'espèces généralistes, peut avoir des conséquences importantes en termes de maintien de la biodiversité, et de bio-indication. La prise en compte en bio-indication d'espèces principalement structurées par des processus de dispersion tendrait en effet à masquer l'effet de l'environnement local sur les communautés entraînant une mauvaise évaluation de la qualité de l'eau. Dans cette perspective, l'effet de masse est particulièrement pointé du doigt puisqu'il permet

à une espèce de sortir de sa niche, uniquement grâce à l'apport d'individus depuis des populations sources. Les espèces soumises à ce mécanisme ne devraient donc pas être prises en compte dans les méthodes de bio-indication se basant sur les diatomées. Or ce paramètre n'est aujourd'hui pas considéré, en partie à cause du manque de connaissances sur le sujet.

Pour les disperseurs actifs, tels que les macro-invertébrés, la mesure de traits spécifiques signant les capacités de dispersion des espèces couplée à la détection de signaux spatiaux peut être utilisée pour détecter ce type de processus. Pour les disperseurs passifs, cela n'est actuellement pas possible et seules des preuves indirectes (comme par exemple, des corrélations négatives entre distances spatiales et valeurs de diversité  $\beta$ ) sont utilisées pour mettre en lumière ces processus (Bottin et al., 2016; Goldenberg Vilar et al., 2014; Jamoneau et al., 2018), ce qui fait que l'influence de l'effet de masse n'a jamais été réellement quantifiée sur les communautés de diatomées. Nous sommes ici confrontés à un verrou scientifique qui rend difficile la compréhension de l'importance relative des conditions environnementales et de l'effet de masse dans la détermination des niches véritablement réalisées par les espèces soumises à ces processus.

Ce processus spatial qui façonne les communautés diatomiques dépend beaucoup de la structure physique du paysage. Les réseaux dendritiques comme les rivières contraignent fortement la dispersion des disperseurs passifs uniquement à travers des trajectoires prédéfinies (Brown et al., 2011; Grant et al., 2007; Liu et al., 2013). Chercher à mesurer une structure spatiale des communautés en se basant sur des distances terrestres non restreintes, dont l'utilisation est pourtant très répandue, n'est donc pas adapté.

Dans ce contexte, la problématique de ce mémoire peut ainsi se résumer en une triple interrogation :

- Quelle est l'influence relative des processus déterministes, comme le *species sorting*, et des processus stochastiques, comme les processus de dispersion, sur la structuration des communautés de diatomées en fonction de l'échelle spatiale considérée ?
- Au sein des processus de dispersion, comment est-il possible d'isoler et de clairement mettre en évidence le phénomène d'effet de masse ?
- Et quelles sont les implications de cette éventuelle mise en évidence de l'effet de masse sur l'utilisation de communautés biotiques en tant qu'outils de bio-indication de la qualité de l'eau ?





## **2 Méthodologie : description des bases de données et des principales méthodes utilisées**

Ce chapitre vise à présenter les bases de données utilisées au cours de cette thèse ainsi que les principales méthodes utilisées. Les détails propres à chaque méthode ainsi que tous les tests statistiques réalisés seront détaillés dans chaque chapitre.

### **2.1 Bases de données**

#### **2.1.1 Base de données Pandore**

La Directive Cadre européenne sur l'Eau du 23 octobre 2000 établit un cadre pour une politique communautaire dans le domaine de l'eau. Cette politique a pour enjeux la protection et la gestion de l'ensemble des milieux aquatiques. Afin d'atteindre cet objectif, il est nécessaire de passer par une étape de diagnostic de l'état écologique des masses d'eau. Pour cela, plusieurs bio-indicateurs avaient été développés en France bien avant la mise en place de la Directive. Toutefois, ces indices bien qu'efficaces ne répondaient pas à toutes les exigences de la DCE et de nouveaux indices ont dû être développés à l'échelle de l'ensemble du territoire. Or, à cette époque, il n'existait pas de base de données hydrobiologiques centralisée au niveau national qui regroupe les résultats d'évaluation des milieux par les indices. C'est dans ce contexte qu'est née la base Pandore constituée par l'Irstea en 2003. Elle devait répondre aux besoins techniques et scientifiques induits par la DCE et a été alimentée par des données produites par les 6 agences de l'eau et les 22 Directions Régionales de l'Environnement, de l'Aménagement et du Logement (DREAL). Ces données couvrent différents maillons biologiques (poissons, macro-invertébrés, macrophytes, et diatomées) ainsi que des données physico-chimiques.

## Diatomées

La forme des données biologiques concernant les diatomées présentes dans la base est intrinsèquement liée à l'utilisation première des données : la bio-indication. Les données présentes ont toutes été récoltées selon le protocole de la norme NFT 90-354 (AFNOR, 2007). Celle-ci prévoit que l'échantillonnage des diatomées doit être effectué dans la zone lotique de la station échantillonnée, sur des supports immergés, durs et stables, si possible naturels. Au moins 5 supports distincts doivent être échantillonnés jusqu'à atteindre une surface minimale de  $100\text{ cm}^2$ , en évitant les zones ombragées. Une fois l'échantillon préparé pour l'observation au microscope, le comptage et l'identification porte sur 400 unités diatomiques au minimum. Les diatomées cassées ne sont prises en compte que si au moins les 3/4 du frustule sont visibles et identifiables. Les données obtenues sont donc sous la forme d'abondances relatives dont le total avoisine les 400 individus. Les données utilisées au cours de cette thèse couvrent une période allant de 2007 à 2013 comprenant environ 1 500 taxons (espèces, sous-espèces, variétés ou formes), ce chiffre variant avec l'évolution de la nomenclature, et plus de 10 000 relevés diatomiques répartis sur plus de 4 000 stations géo-référencées. Afin de limiter les fluctuations de communautés liées à des variations saisonnières, seuls les échantillons prélevés entre les mois de mai et d'octobre ont été conservés. En raison de la diversité des origines de ces données (voir le paragraphe 2.1.1), celles-ci présentent une forte hétérogénéité. Une première étape d'homogénéisation taxonomique avant toute analyse est donc nécessaire et a été réalisée à l'aide du logiciel Omnidia 5.3 (Lecointe et al., 1993) et de diverses bases de données en ligne (<http://www.algaebase.org/>; <https://westerndiatoms.colorado.edu/>).

## Physico-chimie

Des mesures physico-chimiques sont régulièrement réalisées sur ces même stations. Celles-ci comprennent le pH, la conductivité (mesurée à  $20\text{ °C}$ ,  $\text{mS cm}^{-1}$ ), la concentration en oxygène dissous (%), la demande biologique en oxygène ( $\text{mg de } {}_8\text{O}$  consommés par litre en 5 jours d'incubation à  $20\text{ °C}$ ), et les concentrations en phosphore total ( $\text{mg L}^{-1}$  de  ${}_{15}\text{P}$ ), en orthophosphates ( $\text{mg L}^{-1}$  de  ${}_{15}\text{P}$ ), en ammonium ( $\text{mg L}^{-1}$  de  ${}_7\text{N}$ ), en nitrates ( $\text{mg L}^{-1}$  de  ${}_7\text{N}$ ) et en nitrites ( $\text{mg L}^{-1}$  de  ${}_7\text{N}$ ). Une moyenne de ces paramètres a été calculée pour chaque station sur une période de 45 jours (30 jours avant le prélèvement diatomique et 15 jours après ce prélèvement).

### 2.1.2 Base de données du NAWQA

Les données américaines utilisées au cours de cette thèse proviennent de la base de données du *National Water-Quality Assessment Program*. 1078 stations de cette base échantillonnées entre 1993 et 2011 ont été utilisées au cours de cette thèse.

#### Diatomées

Les diatomées de cette base ont été échantillonnées suivant un protocole standardisé proche de celui de l'IBD, la principale différence résidant dans le nombre minimal de frustules identifiés (600 aux États-Unis contre 400 pour l'IBD). Comme pour la base de données Pandore, l'ensemble des échantillons ont été prélevés entre les mois de mai et d'octobre afin limiter les fluctuations liées à des variations saisonnières. Cette base souffre du même problème d'hétérogénéité que la base Pandore puisque sa construction a nécessité la collaboration d'un grand nombre d'opérateurs. De la même manière, les données taxonomiques ont donc dû être homogénéisées à l'aide des mêmes outils.

#### Physico-chimie

Les mesures de physico-chimie présentes dans cette base de données sont les mêmes que celles présentées précédemment. Une moyenne de ces paramètres a été, ici aussi, calculée pour chaque station sur une période de 45 jours (30 jours avant le prélèvement diatomique et 15 jours après ce prélèvement).

### 2.1.3 Paramètres climatiques

Afin d'estimer les conditions climatiques de chaque station, en France et aux US, nous avons utilisé les prédictions du modèle WorldClim (Hijmans et al., 2005). Ce modèle présente un ensemble de couches climatiques globales avec une résolution spatiale de l'ordre du km<sup>2</sup>. Ces données climatiques incluent des estimations annuelles (précipitations annuelles, températures moyennes annuelles), des estimations des variations saisonnières (écarts de températures et de précipitations) et des estimations des conditions extrêmes (températures moyennes des mois les plus froids/chauds, précipitations pendant les saisons les plus humides/sèches).

## 2.1.4 Autres paramètres

En plus des paramètres physico-chimiques et climatiques déjà évoqués, l'altitude (m) et la pente (%) ont été mesurées sur chaque station avec une précision d'une seconde d'arc grâce aux données récoltées par la *Shuttle Radar Topography Mission* (Farr et al., 2007).

## 2.2 Principales méthodes utilisées

Cette section a pour objectif d'introduire les principales méthodes utilisées au cours de cette thèse. Les précisions méthodologiques propres à chacune des problématiques seront apportées dans chacun des chapitres.

### 2.2.1 Mesures de la biodiversité

La notion de diversité tient un rôle central en écologie des communautés. Elle a pour but de mesurer la variabilité des, et entre, les communautés biologiques et permet d'appréhender les mécanismes sous-jacents à la structuration de ces communautés. Elle est classiquement séparée en trois composantes complémentaires et interdépendantes que sont les diversités  $\alpha$ ,  $\beta$  et  $\gamma$  (Whittaker, 1972).

#### Diversité $\alpha$

La diversité  $\alpha$  correspond à la diversité mesurée en un lieu donné, à l'échelle de la communauté. Dans sa forme la plus simple, elle correspond à la richesse spécifique de la communauté observée. Cette mesure ne rend en aucun cas compte de l'équitabilité (*evenness*) de la communauté, et une même valeur de diversité  $\alpha$  peut cacher des réalités biologiques très différentes. La diversité  $\alpha$  est souvent considérée comme la résultante des processus déterministes qui agissent à l'échelle locale, et en particulier au *species sorting*. Cependant, l'effet de processus, notamment stochastiques, à plus large échelle n'est pas à sous estimer puisque les processus locaux ne peuvent sélectionner que des espèces présentes dans le *pool* régional, dont la composition est en partie soumise à des processus stochastiques.

## Diversité $\gamma$

A l'instar de la diversité  $\alpha$ , la diversité  $\gamma$  est une mesure de richesse spécifique, mais cette fois à une plus large échelle englobant un ensemble d'habitats. La notion de *pool* régional d'espèces est régulièrement assimilée au terme de diversité  $\gamma$  bien qu'il n'existe pas d'égalité stricte entre les deux, le *pool* régional d'espèces étant théoriquement la somme de toutes les espèces dans une région donnée. La diversité  $\gamma$  est régulièrement associée aux processus stochastiques puisqu'elle est le reflet des processus de spéciation, d'extinction, ou encore de dispersion qui agissent à l'échelle régionale.

## Diversité $\beta$

La diversité  $\beta$  permet de faire le lien entre les deux mesures de diversité précédemment évoquées et traduit ainsi la notion de dissimilarité (ou parfois de similarité) entre deux communautés. Numériquement, elle a été historiquement définie comme la simple valeur scalaire des diversités  $\alpha$  et  $\gamma$  (équation 2.1) par Whittaker (1960; 1972).

$$\beta = \frac{\gamma}{\alpha} \quad (2.1)$$

Les méthodes de calcul ont depuis beaucoup évolué en de nombreuses définitions mathématiques offrant de nombreux et puissants outils pour étudier les différents aspects de ce concept de variabilité spécifique (Socolar et al., 2016; Tuomisto, 2010a,b). Une des avancées majeures en la matière a été apportée par Baselga (2010) qui a proposé une mesure de la diversité  $\beta$  qui ne soit plus limitée à une comparaison de sites deux à deux. Cette mesure multi-sites a l'intérêt majeur de permettre les mesures de la diversité  $\beta$  à l'échelle régionale, et ainsi de mieux rendre compte des processus agissant à cette échelle. C'est cette définition de la diversité  $\beta$  que nous avons choisi d'utiliser au cours de cette thèse (équation 2.2).

$$\beta_{SOR} = \frac{[\sum_{i < j} \min(b_{ij}, b_{ji})] + [\sum_{i < j} \max(b_{ij}, b_{ji})]}{2 \cdot [\sum_i S_i - S_T] + [\sum_{i < j} \min(b_{ij}, b_{ji})] + [\sum_{i < j} \max(b_{ij}, b_{ji})]} \quad (2.2)$$

Où  $S_i$  est le nombre d'espèces présentes dans le site  $i$ ,  $S_T$  est le nombre d'espèces présentes dans l'ensemble des sites, et  $b_{ij}$  et  $b_{ji}$  les nombres d'espèces exclusives aux sites  $i$  et  $j$ .

Dans la même étude, Baselga a également introduit le concept de partition de la diversité  $\beta$  en deux composantes, le *turnover* et la *nestedness*. La première correspond aux remplacements d'espèces d'un site à un autre le long de gradients environnementaux alors que la seconde composante permet de refléter des processus d'emboîtement entre différentes communautés (Fig. 2.1)(Baselga, 2010).

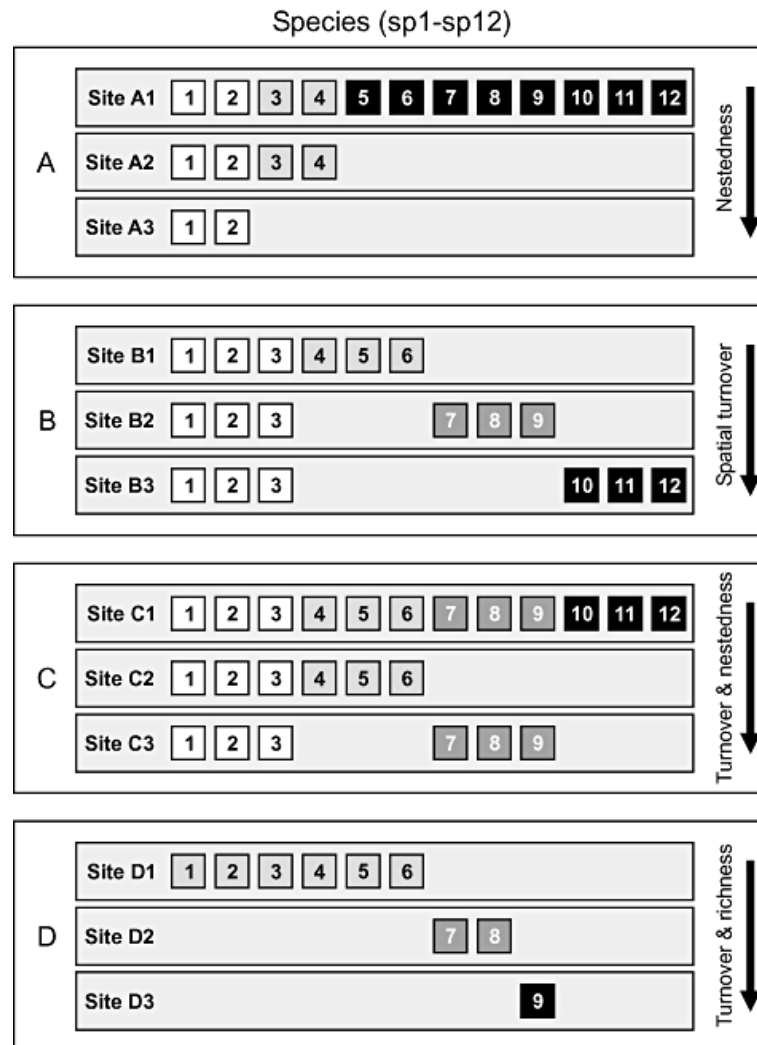


FIGURE 2.1 – Exemples théoriques pour illustrer les notions de *turnover* et de *nestedness* par Baselga (2010).

L'utilisation de la diversité  $\beta$  représente un apport considérable pour déterminer l'importance et la nature des processus déterministes (liés au concept de niche écologique) et stochastiques (liés aux processus de dispersion) sur la structuration des communautés le long de gradients environnementaux (Chase and Myers, 2011). Cependant, les études s'intéressant aux contributions relatives sur la structuration des communautés de

ces deux types de mécanismes restent relativement peu nombreuses, en particulier dans un cadre multi-échelles explicite.

### **2.2.2 Cadre multi-échelles explicite : développement de la méthode des fenêtres paysagères**

Un des enjeux de cette thèse était d'étudier l'influence des différents processus dans un cadre spatial explicite, permettant d'étudier l'évolution de ces influences le long de gradients spatiaux. Nous avons choisi de développer une méthode, basée sur la création de fenêtres paysagères, nous permettant d'en assurer la répliquabilité, et de nous affranchir des limites imposées par les bassins versants ou les barrières biogéographiques naturelles, qui sont souvent des limites importantes pour ce type d'approche.

Les fenêtres ont été pensées comme des polygones carrés, placés aléatoirement ou à distance régulière selon les besoins de l'étude et les contraintes de la zone d'étude, dont la taille peut varier de quelques mètres à plusieurs milliers de kilomètres. Pour être conservées dans l'analyse, les fenêtres doivent contenir un nombre minimum de sites, qui une fois encore est personnalisable, et ces sites doivent être distribués sur l'ensemble du polygone afin de s'assurer que l'échantillonnage réalisé correspond bien à l'échelle spatiale que la fenêtre souhaite décrire. Le système de projection de Robinson a été retenu pour cette technique puisque la distorsion engendrée par cette projection est très faible assurant ainsi une équitabilité entre les surfaces des fenêtres d'une même échelle lors d'études à grande échelle, comme c'est le cas ici avec l'utilisation de données françaises et américaines. Ce développement méthodologique a été conduit en langage R (R Core Team, 2018) et le script est disponible librement sur internet (<https://github.com/tleboucher/landscape.windows>).

Lors de notre travail, cette technique a été utilisée pour appréhender l'évolution de la diversité  $\beta$  en fonction des échelles spatiales et en fonction d'un gradient de nutriments, et pour mesurer l'influence des processus déterministes et stochastiques sur la structuration des communautés de diatomées, en France et aux États-Unis, étude présentée dans le chapitre 3. Lors de cette étude, une grille de 50 km par 50 km a été utilisée pour placer le centre de chaque fenêtre de manière régulière résultant en 9006 fenêtres potentielles par échelle. Lors de cette étude, nous avons fait varier la taille des fenêtres entre 200 km et 2000 km. Afin d'être validées, les fenêtres devaient contenir au moins 25 sites pour chacun des niveaux d'enrichissement en nutriments, contrainte imposée pour les mesures de diversité  $\beta$  qui nécessitent ce nombre de 25 sites pour obtenir une mesure fiable de la diversité  $\beta$  à l'échelle régionale (Bennett et al., 2010). Une grille de 3 cases par 3 cases a été utilisée pour vérifier la bonne distribution des sites dans chaque fenêtre, et ainsi s'assurer qu'ils soient bien représentatifs de l'ensemble de la fenêtre, et donc de l'échelle spatiale associée. Pour chaque fenêtre, 80 % des cases devaient contenir au moins un

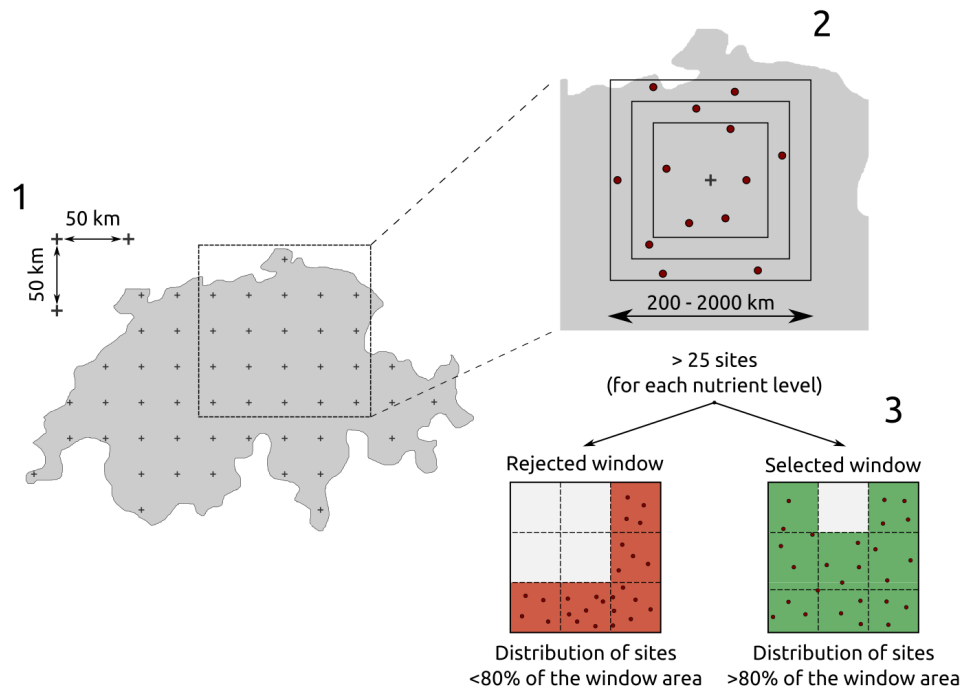


FIGURE 2.2 – Représentation schématique de l'approche par fenêtres paysagères utilisées lors de l'étude présentée dans le chapitre 3. 1 - Le centre de chaque fenêtre a été placé régulièrement tous les 50 km le long de transect latitudinaux et longitudinaux. 2 - De fenêtres mesurant entre 200 km et 2000 km ont été créées et devaient contenir au minimum 75 sites (25 pour chaque niveau de nutriments). 3 - Ces sites devaient, pour chaque niveau de nutriments, être distribués sur plus de 80 % de la surface de la fenêtre (utilisation d'une grille de 3 cases par 3 cases).

site pour que celle-ci soit validée. Cette procédure a été schématiquement résumée sur la figure 2.2.

### 2.2.3 Étude de l'influence relative des processus déterministes et stochastiques

L'intérêt des écologues pour les modèles nuls a émergé aux cours de la dernière décennie en particulier avec les travaux de Kraft, Chase et Myers et de leurs collaborateurs (Chase et al., 2011; Chase and Myers, 2011; Kraft et al., 2011; Myers et al., 2013). L'enjeu est de pouvoir démêler l'influence des processus qui considèrent que les variations de la diversité  $\beta$  le long de gradient environnementaux sont prédictibles et liés à des processus déterministes locaux (avec donc un fort impact sur la diversité  $\alpha$ ), de l'influence

des processus stochastiques qui prédisent que la diversité  $\beta$ , au contraire, ne change pas systématiquement le long de ces mêmes gradients environnementaux et reste liée à des processus de dispersion, aux taux de spéciation, de colonisation des espèces, etc (avec donc un fort impact sur la diversité  $\gamma$ ). Démêler l'influence de ces deux types de processus sur la diversité  $\beta$  n'est pourtant pas chose aisée puisque par définition, la diversité  $\beta$  représente le lien qui unit les échelles locale (diversité  $\alpha$ ) et régionale (diversité  $\gamma$ ).

L'approche proposée par les modèles nuls est de s'intéresser à comment, et de combien, la diversité  $\beta$  observée dans une région donnée s'éloigne de ce qui est attendu d'une situation où il est assumé que les processus déterministes sont sans importance et sans effet sur la structuration des communautés. La forte déviation entre les valeurs mesurées et les valeurs prédites par ce procédé met alors en exergue l'importance des processus déterministes sur la structuration des communautés, alors qu'au contraire, lorsque le modèle est proche de la réalité et que la déviation entre les valeurs mesurées et les valeurs prédites est faible, cela traduit une prévalence des processus stochastiques. Une précaution doit cependant être prise lorsque de tels modèles sont utilisés pour comparer des communautés de différentes régions biogéographiques (différents bassins versants, différents climats, etc.). En effet, ces régions risquent de différer en ce qui concerne la taille du *pool* d'espèces, ce qui risque d'influencer le résultat des modèles nuls ne prenant pas compte cette variation. Une approche contrôlant la taille de ce *pool* d'espèces, estimé par les valeurs de diversité  $\gamma$  est alors nécessaire. C'est ce type d'approche que nous avons choisi d'utiliser lors de cette thèse en adaptant un modèle déjà existant (voir Kraft et al., 2011) à la mesure de diversité  $\beta$  multi-sites.

La procédure retenue redistribue les individus d'une communauté de manière aléatoire, assumant ainsi l'absence d'impact des processus déterministes, mais en conservant la même diversité  $\gamma$  dans la communauté "simulée" et dans la communauté "observée" afin de tenir compte de la variation en taille des *pool* d'espèces entre les régions. Les nouveaux relevés floristiques ainsi obtenus sont alors convertis en matrices indiquant la présence ou l'absence des espèces dans chacun des sites afin de calculer la diversité  $\beta$  dans ces communautés "simulées" exactement de la même manière que pour les communautés réellement observées (suivant l'équation 2.2 définie précédemment). Afin d'obtenir la valeur la plus aléatoire possible, cette opération est répétée 999 fois afin d'obtenir 1000 valeurs théoriques pour chaque valeur de diversité  $\beta$  réellement observée. Les déviations entre les deux valeurs sont alors calculées de la façon suivante :

$$\beta_{dev} = \frac{\beta_{obs} - \beta_{null}}{sd_{null}} \quad (2.3)$$

Où  $\beta_{obs}$  est la valeur originelle (observée) de diversité  $\beta$ ,  $\beta_{null}$  la moyenne des 1000 valeurs de diversité  $\beta$  simulée et  $sd_{null}$  l'écart-type des 1000 valeurs de diversité  $\beta$  simulée.

Avec un tel modèle, les valeurs de déviation ne sont pas bornées et peuvent varier de  $-inf$  à  $+inf$ . Comme indiqué précédemment, de fortes valeurs de déviation indiqueront une importance des processus déterministes sur la structuration des communautés de diatomées, alors que des valeurs faibles ou négatives indiqueront au contraire une forte structuration des communautés par des processus stochastiques.

Malgré leur grand intérêt et la possibilité qu'ils offrent d'étudier les patrons spatiaux de la diversité  $\beta$ , ces modèles nuls sont encore assez peu utilisés dans des cadres multi-échelles bien définis. Ils souffrent également d'une critique d'une partie de la communauté scientifique. Ulrich et al. (2017) défendent l'idée que le contrôle de la diversité  $\gamma$  n'est pas suffisant et conduit dans certains cas à l'échec du rejet de l'hypothèse nulle, c'est-à-dire l'absence total d'influence des processus déterministes sur la structuration des communautés, conduisant ainsi à une conclusion erronée (erreur de Type II). C'est pour cette raison et pour éviter ces critiques que nous avons inclus cette approche dans le cadre offert par la méthode des fenêtres paysagères qui permet de mieux contrôler l'effet du *pool* d'espèces (diversité  $\gamma$ ) dans le modèle.

## 2.2.4 Influence des processus de dispersion et modélisation de la distribution des espèces

Dans le cadre du chapitre 4 visant à démêler l'influence des processus de dispersion (et plus précisément celle de l'effet de masse), de celle des processus liés à la niche, à l'échelle de l'espèce, des approches innovantes ont été utilisées et combinées : les distance-based Moran's Eigenvectors Maps (dbMEM) et les Boosted Regressions Trees (BRT).

### dbMEM

Les dbMEM sont une famille d'analyses visant à mieux modéliser l'influence des processus spatiaux, et donc des processus de dispersion, sur les communautés. Elles dérivent des Principal Coordinates of Neighbour Matrices (PCNM) développées par Borcard and Legendre au début des années 2000 (2002) et permettent la création d'*eigenvectors* décrivant les relations spatiales (connectivité) entre les sites étudiés (Fig. 2.3). Ces *eigenvectors* sont alors utilisés comme prédicteurs spatiaux dans des régressions linéaires, des analyses canoniques, etc, et permettent de quantifier les patrons purement géographiques. Toutefois, la présence de signaux spatiaux structurant les communautés peut parfois être le simple résultat de la présence de gradients environnementaux spatialisés non mesurés, même si la prise en compte de très nombreux gradients environnementaux peut permettre de limiter cet effet.

Les processus spatiaux dépendent également beaucoup de la structure physique du paysage. Les réseaux dendritiques comme les rivières contraignent fortement la dispersion des disperseurs passifs uniquement à travers des trajectoires prédéfinies (Brown et al., 2011; Grant et al., 2007; Liu et al., 2013). Chercher à mesurer une structure spatiale des communautés en ce basant sur des distances terrestres non restreintes, dont l'utilisation est pourtant très répandue, n'est donc pas adapté. Les Asymmetric Eigenvector Maps (AEM) (Blanchet et al., 2008a) permettent de prendre en compte la structuration contrainte des réseaux de rivières et la relation asymétrique causée par le courant qui existe entre les sites. Les AEM sont en effet une forme de dbMEM permettant de modéliser l'influence du spatial en présence d'un processus physique asymétrique (Fig. 2.4). Les *eigenvectors* ainsi obtenus ont été utilisés comme variable explicative pour essayer de prédire la présence ou l'absence des espèces et mettre en évidence le rôle des processus de dispersion dans la structuration des communautés.

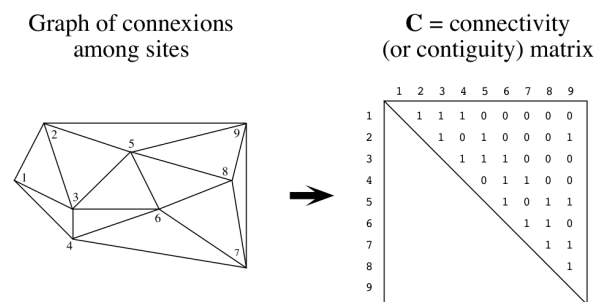


FIGURE 2.3 – Graphe de connexions entre les sites étudiés en une matrice binaire de connectivité (d'après P. Legendre).

## Patrons de co-occurrences

Déjà utilisés par Bottin et al. (2016) pour mettre en évidence le rôle de la dispersion dans la structuration des communautés, les processus de dispersion ont des effets variés sur les patrons de co-occurrences. Puisque l'effet de masse peut permettre aux espèces de persister dans des habitats *a priori* peu favorables (c'est-à-dire, loin de leur optimum écologique), il tend à augmenter le chevauchement des niches par rapport aux espèces dont les processus sont principalement axés sur les niches. En revanche, la limitation de la dispersion tend à réduire le chevauchement des niches en empêchant les espèces d'atteindre tous les habitats appropriés. Ces patrons de chevauchement des niches peuvent être déduit par une analyse des co-occurrences des espèces. Ainsi, de fortes associations négatives peuvent suggérer la prévalence de processus de niche ou de la dispersion limitée sur la distribution des espèces, tandis que de faibles associations négatives, combinées à des signaux spatiaux, indiquent une forte influence de l'effet de masse sur la distribution des espèces.



## BRT

Une fois les prédicteurs spatiaux de l'effet de masse obtenus, nous avons opté pour l'utilisation d'une technique de modélisation de la distribution des espèces afin de mesurer l'influence relative de l'effet de masse et de l'environnement sur la structuration de chaque espèce. Les BRTs, le modèle retenu, sont la combinaison de deux techniques : les algorithmes d'arbres de décision et les méthodes de *boosting*. De la même façon que pour les *Random Forest*, les BRTs génèrent de façon répétée de nombreux arbres décisionnels pour améliorer la précision du modèle. La principale différence avec les *Random Forest* est qu'après l'ajustement du premier arbre, le modèle tiendra compte de l'erreur dans la prédiction de cet arbre pour l'arbre suivant, et ainsi de suite. En tenant compte de l'ajustement des arbres précédents qui sont construits, le modèle essaie continuellement d'améliorer sa précision. Cette approche séquentielle est unique au *boosting* et a déjà donné de bons résultats lorsqu'ils ont été appliqués sur des micro-organismes (Pajunen et al., 2016). Les détails concernant le paramétrage de ces modèles sont indiqués dans le chapitre 4.



# 3 Influences des processus déterministes et stochastiques sur la structuration des communautés de diatomées benthiques : approche multi-échelle et impact des perturbations humaines

Ce chapitre présente un article accepté dans la revue *Journal of Biogeography*.

## 3.1 Contexte et objectifs de l'étude

Dans le contexte actuel de déclin rapide de la biodiversité, mieux comprendre la contribution des différents processus, déterministes et stochastiques, agissant sur la structuration des communautés biologiques est, plus que jamais, un enjeu majeur pour assurer le suivi et la conservation de la biodiversité. Une bonne compréhension de ces mécanismes est également primordiale en bio-indication puisque la pertinence des indices repose sur notre capacité à distinguer les variations dues à l'influence de processus locaux d'origine naturelle ou non, de celles dues à des processus régionaux, comme les processus de dispersion. Nos travaux ont donc cherché à évaluer la part relative de ces contributions sur la structuration des communautés de diatomées benthiques, à différents niveaux de pressions d'origine anthropique (charge en nutriments).

L'évolution et l'apparition de nouveaux indices de diversité  $\beta$  offre aujourd'hui de puissants outils pour explorer ce type de questions (Socolar et al., 2016) induisant ainsi un regain d'intérêt pour les mesures de diversité  $\beta$ . Ainsi, la diversité  $\beta$  est reconnue pour être le résultat de l'interaction de gradients biogéographiques, environnementaux et spatiaux qui interagissent pour créer une structuration en partie déterministe et en partie stochastique des *pools* d'espèces locaux et régionaux (Tonkin et al., 2015; Viana et al.,

2016). Le *species sorting*, littéralement tri des espèces, résultat des filtres environnementaux et des interactions biotiques, ainsi que les processus de dispersion apparaissent ainsi comme les deux moteurs principaux de la structuration de la diversité  $\beta$  (Chase and Leibold, 2003; Soininen, 2007), mais leur influence est toutefois contrainte par l'effet de l'échelle spatiale considérée (locale ou régionale). Les approches multi-échelles sont donc vivement recommandées (Huston, 1994; Mac Nally et al., 2004) pour mieux comprendre comment les processus locaux et régionaux affectent la biodiversité, bien que peu d'études les aient réellement mises en œuvre. Ceci est l'une des originalités de notre étude puisque nous avons voulu palier cette absence en nous intéressant, pour la première fois à notre connaissance, aux patrons de diversité  $\beta$  de communautés soumises à différents niveaux d'enrichissement en nutriments, à travers une approche multi-échelles.

## 3.2 Principaux résultats

Grâce à une large base de données intercontinentales incluant 2212 stations d'échantillonnage en France et 1078 aux États-Unis, collectées entre 1993 et 2011, et à une technique innovante de fenêtres paysagères, nous avons pu obtenir les résultats suivants.

Premièrement, nous avons pu confirmer que la diversité  $\beta$  augmentait avec l'échelle spatiale considérée, mais que cette tendance était clairement modulée par l'enrichissement en nutriments, l'enrichissement en nutriment étant associé à un déclin global de la diversité  $\beta$ , notamment en raison de l'élimination des espèces spécialistes et par leur remplacement par des espèces généralistes dans les communautés.

Deuxièmement, nous avons pu montrer que les processus d'assemblage locaux étaient le principal moteur de structuration des communautés de diatomées, indifféremment en France et aux États-Unis, et que leurs influences sur la diversité  $\beta$  ne faisaient que s'accroître avec l'échelle spatiale, au détriment des processus structurant la diversité  $\gamma$ . Cependant, l'enrichissement, ici aussi, a joué un rôle prépondérant modulant la contribution relative du *species sorting* au profit des processus de dispersion, tout au long du gradient spatial étudié.

# Diatom $\beta$ -diversity in streams increases with spatial scale and decreases with nutrient enrichment across regional to sub-continental scales

Thibault Leboucher<sup>1+\*</sup>, William R. Budnick<sup>2+</sup>, Sophia I. Passy<sup>2</sup>, Sébastien Boutry<sup>1</sup>, Aurélien Jamoneau<sup>1</sup>, Janne Soininen<sup>3</sup>, Wim Vyverman<sup>4</sup> and Juliette Tison-Rosebery<sup>1</sup>

<sup>1</sup> Aquatic Ecosystems and Global Changes Research Unit, IRSTEA, 50 avenue de Verdun, 33612 Cestas, France

<sup>2</sup> Department of Biology, University of Texas at Arlington, Box 19498, Arlington, TX 76019, USA

<sup>3</sup> Department of Geosciences and Geography, University of Helsinki, PO Box 64, Helsinki, FIN-00014, Finland

<sup>4</sup> Ghent University, Biology Department, Protistology and Aquatic Ecology, Krijgslaan 281 S8, 9000 Gent, Belgium

\* Corresponding author : thibault.leboucher@irstea.fr

+ Co-first author

## Abstract

**Aim** To quantify the relative contributions of local community assembly processes vs.  $\gamma$ -diversity in driving  $\beta$ -diversity, and to assess how spatial scale and anthropogenic disturbance (*i.e.* nutrient enrichment) interact to dictate which driver dominates.

**Location** France and the United States.

**Time period** 1993-2011

**Major taxa studied** Freshwater stream diatoms

**Methods**  $\beta$ -diversity along a nutrient enrichment gradient was examined across multiple spatial scales.  $\beta$ -diversity was estimated using multi-site Sørensen dissimilarity. We estimated the relative importance of specialists vs. generalists using Friedley coefficient, and

the contribution of local community assembly vs.  $\gamma$ -diversity on  $\beta$ -diversity across spatial scales, with a null model. Finally, we estimated the response of  $\beta$ -diversity to environmental and spatial factors by testing the correlations between community, environmental and geographical distance matrices with partial Mantel tests.

**Results**  $\beta$ -diversity generally increased with spatial scales but the rate of increase depended on nutrient enrichment level.  $\beta$ -diversity decreased significantly with increasing nutrient enrichment level due to the loss of specialist species. Local assembly was an important driver of  $\beta$ -diversity especially under low nutrient enrichment. Significant partial Mantel correlations were observed between diatom  $\beta$ -diversity and pure environmental distances under these conditions, highlighting the role of species sorting in local assembly processes. Conversely, in heavily enriched sites, only spatial distances were significantly correlated with  $\beta$ -diversity, which indicated a substantial role of dispersal processes.

**Main conclusions** Nutrient supply mediated the expected increase of  $\beta$ -diversity with spatial scales. Across spatial scales,  $\beta$ -diversity was more influenced by local assembly processes rather than by  $\gamma$ -diversity. Nutrient enrichment was associated with an overall decline in diatom  $\beta$ -diversity and a shift in assembly processes from species sorting to dispersal, notably due to the elimination of some specialists and their subsequent replacement by generalists.

**Keywords :**  $\beta$ -diversity, community assembly, diatom communities, dispersal processes,  $\gamma$ -diversity, landscape windows approach, multi-scale approach, null model, species sorting

## Introduction

Perhaps one of the most interesting facets in the study of biodiversity is the consistent shift in species membership as one samples biological communities across space and time. The pattern of variation in local community composition ( $\alpha$ -diversity), given a regional species pool ( $\gamma$ -diversity), was dubbed  $\beta$ -diversity by R.H. Whittaker more than fifty years ago (Whittaker, 1960, 1972), and was originally defined in his work as “the extent of change in community composition in relation to a complex-gradient of environment, or a pattern of environments”. The concept of  $\beta$ -diversity has since morphed from a simple scalar value (*i.e.*  $\beta = \gamma/\bar{\alpha}$ ), into numerous mathematical definitions, which now include similarity/distance and variance-based measures (Tuomisto, 2010a,b; Tuomisto and Ruokolainen, 2006). The preponderance of new quantitative  $\beta$ -diversity estimators/indices and conceptualizations has thus given researchers powerful tools to study different aspects of Whittaker’s concept of community variation (Socolar et al., 2016). Consequently,  $\beta$ -diversity has a recent surge of interest because differences in the quantitative forms and shapes of many widely studied macroecological patterns (*e.g.* species-area relationships and species abundance distributions) are associated with variation in  $\beta$ -diversity among different metacommunities and regions.

$\beta$ -diversity is generally caused by an interplay of biogeographic, environmental and spatial gradients interacting to create stochastic and deterministic structuring of the local and regional species pool (Jamoneau et al., 2018; Tonkin et al., 2015; Viana et al., 2016). Species sorting (*i.e.* environmental filtering coupled with biological interactions), and dispersal processes are particularly important drivers of  $\beta$ -diversity (Chase and Leibold, 2003; Soininen, 2007), however, their relative influence are constrained by the effects of spatial scale (local vs. regional). Scale effects on measures of  $\beta$ -diversity, and their drivers, have been of particular interest among ecologists in the last decade because much recent research has shown that the average distance between sampling units, their spatial configuration, and the spatial extent of the sampling region all confound comparisons of  $\beta$ -diversity between study areas (Conroy and Noon, 1996; Mac Nally et al., 2004; Noss, 1983; Wilson and Shmida, 1984). Quantitative multi-scales approaches are thus recommended to better understand how local and regional processes affect  $\beta$ -diversity (Huston, 1994; Mac Nally et al., 2004), and to also identify the appropriate spatial scales for studying mechanisms that structure communities (Mac Nally et al., 2004; Underwood and Chapman, 1996; Willis and Whittaker, 2002). However, few studies have actually implemented quantitative scale-explicit frameworks to explore such relationships.

Recent literature yet tells us that  $\beta$ -diversity tends to vary with spatial scale through different deterministic and stochastic processes. Species sorting, in particular, shapes community composition along environmental gradients (Lack, 1976; Leibold et al., 2004) and thus along spatial gradient when environmental heterogeneity increases with scale. Where environmental gradients persist at geological time scales (*i.e.* under natural conditions), lo-

cal specialization leads to high regional-level community dissimilarity (Gutiérrez-Cánovas et al., 2013), and thus  $\beta$ -diversity tends to increase with spatial scale. Dispersal processes may also influence community dissimilarity by compositional homogenization (lower  $\beta$ -diversity) or differentiation (higher  $\beta$ -diversity) as a function of the rate of dispersal (Dong et al., 2016; Jamoneau et al., 2018; Soininen, 2007). However, it remains poorly understood how species sorting and dispersal processes interact to structure community variation across scales, given that their relative contributions are clearly dependent on the spatial context in which they are measured.

A further complication in the study of scale relationships of  $\beta$ -diversity drivers, is that  $\beta$ -diversity itself is constrained by variation in  $\alpha$ -diversity and  $\gamma$ -diversity along the same environmental and spatial gradients (Kraft et al., 2011). Consequently, to assess the impact of local assembly processes (*i.e.*  $\alpha$ -diversity pattern) on  $\beta$ -diversity, one must account for the influence of the regional species pool (*i.e.*  $\gamma$ -diversity). Null models have proven to be useful in this regard, and have seen widespread application in the analysis of  $\beta$ -diversity pattern and its drivers (*e.g.* Chase et al., 2011; Kraft et al., 2011; Myers et al., 2013), but have not been widely adopted into quantitative multi-scale frameworks.

Scale-dependent relationships of  $\beta$ -diversity drivers (*i.e.* species sorting, dispersal processes and the regional species pool) are further influenced by anthropogenic disturbance, causing extensive impairment of terrestrial and freshwater ecosystems (Vörösmarty et al., 2010). Anthropogenic disturbances have been associated with taxonomic homogenization (loss of  $\beta$ -diversity), but it is unclear which are the key mechanisms and processes contributing to homogenization. One general pattern often reported is the elimination of sensitive species (McKinney and Lockwood, 1999; Olden and Poff, 2003, 2004; Socolar et al., 2016) and subsequent replacement by a few generalist species, usually sharing similar traits (Gutiérrez-Cánovas et al., 2013). Under natural conditions,  $\beta$ -diversity increases with spatial scale owing to increased environmental heterogeneity and habitat specialization (Chase, 2014). However it is unclear if anthropogenic disturbance, by impoverishing the species pool of specialists and increasing the prevalence of generalists, would attenuate the increase in  $\beta$ -diversity with spatial scale.

The replacement of specialist species by generalists is likely the result of differential interactions of anthropogenic disturbance with the underlying deterministic and stochastic processes operating on  $\beta$ -diversity. If anthropogenic disturbances eliminate specialists and promote generalists with broader niches, one consequence is that the effect of species sorting might weaken and species composition would be mostly affected by stochastic processes, including dispersal. However, deterministic and stochastic processes are scale dependent, and it is unknown if anthropogenic disturbance may shift the dominant processes with changing spatial scale.

Our goal was to assess the relative contributions of local community assembly processes vs.  $\gamma$ -diversity, and species sorting vs. stochastic processes, in driving  $\beta$ -diversity, and whether spatial scale and anthropogenic disturbance interact to determine which pro-

cesses dominate. Diatoms represent an appropriate biological model for such research since their communities exhibit biogeographical patterns (Soininen et al., 2016; Vanormelingen et al., 2008; Vyverman et al., 2007), subject to both local and regional processes. We first explored how stream diatom  $\beta$ -diversity responded to different levels of anthropogenic disturbance (*i.e.* mainly agricultural nutrient enrichment) from local to subcontinental scales. We then quantified how  $\beta$ -diversity was driven by local vs.  $\gamma$ -diversity and by spatial vs environmental effects. We tested the following three hypotheses.

1.  $\beta$ -diversity in diatom communities should generally increase with scale due to greater environmental heterogeneity but such increase would be weak in nutrient impacted sites as a result of the loss of sensitive species and impoverished species pool.
2. Under nutrient impacted conditions, a stronger effect of  $\gamma$ -diversity is expected.
3. If anthropogenic disturbances eliminate specialists and promote generalists, species sorting would have a weaker effect on communities under nutrient-impacted conditions and species composition would be mainly constrained by spatial processes, including dispersal.

## Materials and Methods

### Diatoms and environmental datasets

Diatom datasets from France and the United States included a total of 3391 spatially distinct localities (Table 3.1). Samples were collected primarily between May and October in order to reduce seasonal variability in species assemblages. All community and physico-chemical data were gathered in 2011 by the Water Agencies for the French sites and between 1993 and 2011 by the National Water-Quality Assessment (NAWQA) Program for the US sites. In both countries, coordinates for all samples were calculated with WGS84 datum and projected using a Robinson projection (EPSG : 54030). Environmental variables included water pH, specific conductance (at 20 °C, mS cm<sup>-1</sup>), dissolved oxygen saturation (%), concentrations of total phosphorous (mg L<sup>-1</sup> of <sup>15</sup>P), orthophosphate (mg L<sup>-1</sup> of <sup>15</sup>P), ammonium (mg L<sup>-1</sup> of <sup>7</sup>N), nitrate (mg L<sup>-1</sup> of <sup>7</sup>N) and nitrite (mg L<sup>-1</sup> of <sup>7</sup>N). Environmental data were not recorded at the time of community sampling ; therefore, we used the median values of all environmental variable measurements obtained during the 60 days before and the 15 days after the diatom sampling date. We also included geographical variables, *i.e.*, altitude (m, 1 arc-second) and river slope (‰, 1 arc-second), drawn from the Shuttle Radar Topography Mission (Farr et al., 2007) and climatic variables, *i.e.*, annual precipitations (mm) and annual air temperature (°C), drawn from the WorldClim

database (Hijmans et al., 2005). Climate data covered the years 1950–2000 and the average values were included in the analyses.

|               | No.<br>unimpacted<br>sites | No. slightly<br>impacted<br>sites | No. heavily<br>impacted<br>sites | No. species | Years of<br>sampling |
|---------------|----------------------------|-----------------------------------|----------------------------------|-------------|----------------------|
| France        | 636                        | 1163                              | 413                              | 901         | 2011                 |
| United States | 555                        | 358                               | 165                              | 946         | 1993-2011            |

TABLE 3.1 – Details of the stream diatom datasets

In France, diatoms were sampled with a standardized protocol, in line with the European standards (EN 13946, European Commission). In the US, diatoms were sampled following the standardized protocols of the US Geological Survey's National Water Quality Assessment Program (NAWQA). In both countries, cells were identified at 1000 x magnification by examining permanent slides of cleaned diatom frustules (400 valves per slide in France and 600 valves per slide in the US). The French and the US diatom datasets were taxonomically homogenized at the species level using Omnidia 5.3 (Leconte et al., 1993) and different databases (<http://www.algaebase.org/>, <https://western-diatoms.colorado.edu/>).

### Level of nutrient enrichment

To classify the sampling stations along a eutrophication gradient, we first performed principal components analysis (PCA) on power-transformed nutrient concentrations (nitrate, nitrite, ammonia, orthophosphate and total phosphorous). We used Box-cox maximum likelihood estimation (Box and Cox, 1964) to decide the best power-transformation for approximating normality. We then simplified the analysis by retaining only the components with eigenvalues greater than those randomly generated by a broken-stick model (Jackson, 1993; Legendre and Legendre, 2012). Following this procedure, we found that only the first principal component axis (PC1) was significant. PC1 explained 55 % of the total variance and represented a gradient of nutrient enrichment (Fig. 3.1). We used an ascending hierarchical classification (AHC) of the PC1 scores to classify sites into classes of nutrient enrichment based on their Euclidean distances. AHC identified three homogeneous classes optimizing the average silhouette width, *i.e.*, unimpacted, slightly impacted and heavily impacted sites (Table 3.1).

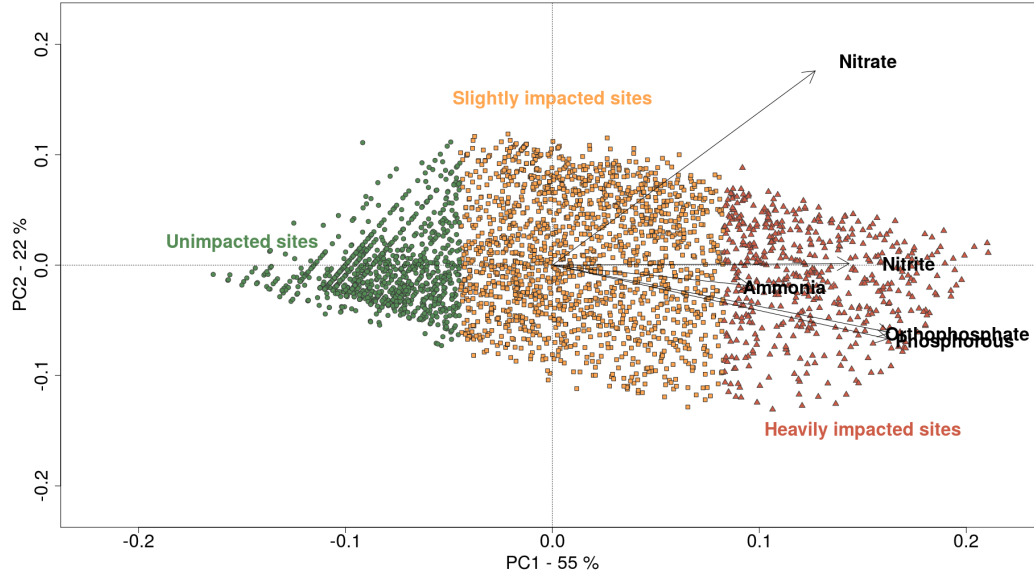


FIGURE 3.1 – Principal components analysis plot for axes 1 and 2 using nitrate, nitrite, ammonia, total phosphorous and orthophosphate concentrations. Environmental variables are shown with arrows and colors represented the nutrient enrichment level.

### Generalist and specialist species

To achieve a classification of species along a generalist-specialist gradient, we implemented the approach by Fridley et al. (2007), which is based on species co-occurrences and allows species characterization without measuring a  $n$ -dimensional niche for each species. This approach derives from a  $\beta$ -diversity statistic and relies on the assumption that generalists should co-occur with many species, whereas specialists should co-occur with relatively few species (Fridley et al., 2007), independently of the species pool size (Zelený, 2009). The continuous metric ( $\theta$ -value or Fridley coefficient) obtained for each species reflects the relative species niche width : low  $\theta$ -values denote a confined niche (specialist species), while high  $\theta$ -values denote a broader niche (generalist species). We measured the mean  $\theta$ -value at each site, weighted by the abundance of each species present at this site, to estimate the relative importance of specialists vs generalists both in France and in the US, following the equation

$$\theta_{site} = \sum_1^n \theta_i \cdot \frac{Ab_i}{Ab_{tot}}$$

where  $n$  is the number of species in the site,  $\theta_i$  is the  $\theta$ -value of species  $i$ ,  $Ab_i$  is the

abundance of species  $i$  in the site and  $Ab_{tot}$  is the total abundance of all species in the site.

## Landscape windows approach

$\beta$ -diversity among sites was examined across multiple spatial scales with square polygons (*i.e.* landscape windows), using R 3.4.2 (R Core Team, 2017). Spatial scales are defined by the size of the landscape windows and varied between  $200 \times 200$  km and  $2000 \times 2000$  km ( $0.4 \times 10^5$  km<sup>2</sup> and  $0.4 \times 10^5$  km<sup>2</sup>, respectively). Latitudinal transects at 50 km intervals were used to place the center of each window. Windows were aligned along these transects at 50 km distances between their centers, resulting in 9006 possible and potentially overlapping windows at each scale (Fig. 3.2). We selected windows that contained at least 25 unimpacted, 25 slightly impacted and 25 heavily impacted sites distributed over at least 80 % of the window's area to ensure equal sampling effort across categories (see Appendix S1). Among the 9006 possible landscape windows, 2514 in both France and the US met our selection criteria. There were no landscape windows below  $6.4 \times 10^5$  km<sup>2</sup> in the US because of the low density of sampling localities. For simplicity, each spatial scale was denoted by the width of its corresponding window, for example the  $200 \times 200$  km ( $0.4 \times 10^5$  km<sup>2</sup>) was denoted as “200 km window”. The creation of landscape windows in R 3.4.2 (R Core Team, 2017) is provided in Appendix S3. In contrast to a simple distance-decay approach, our landscape windows approach explicitly accounts for community dissimilarities and spatial distribution of the sites and thus renders the spatial-scaling pattern of  $\beta$ -diversity more informative, robust and comparable among studies.

## $\beta$ -diversity calculations

We estimated the  $\beta$ -diversity of each window using the multi-site total  $\beta$ -diversity of Baselga-family decomposition of the Sørensen dissimilarity coefficient for species presence-absence data (Baselga, 2010) obtained with the `beta.div.comp` function (R-package : ‘`adespatial`’, Dray et al., 2018) using the equation

$$\beta = \frac{[\sum_{i < j} \min(b_{ij}, b_{ji})] + [\sum_{i < j} \max(b_{ij}, b_{ji})]}{2 \cdot [\sum_i S_i - S_T] + [\sum_{i < j} \min(b_{ij}, b_{ji})] + [\sum_{i < j} \max(b_{ij}, b_{ji})]}$$

where  $S_i$  is the total number of species present in site  $i$ ,  $S_T$  is the total number of species in all sites, and  $b_{ij}$  and  $b_{ji}$  are the number of species exclusive to sites  $i$  and  $j$ .

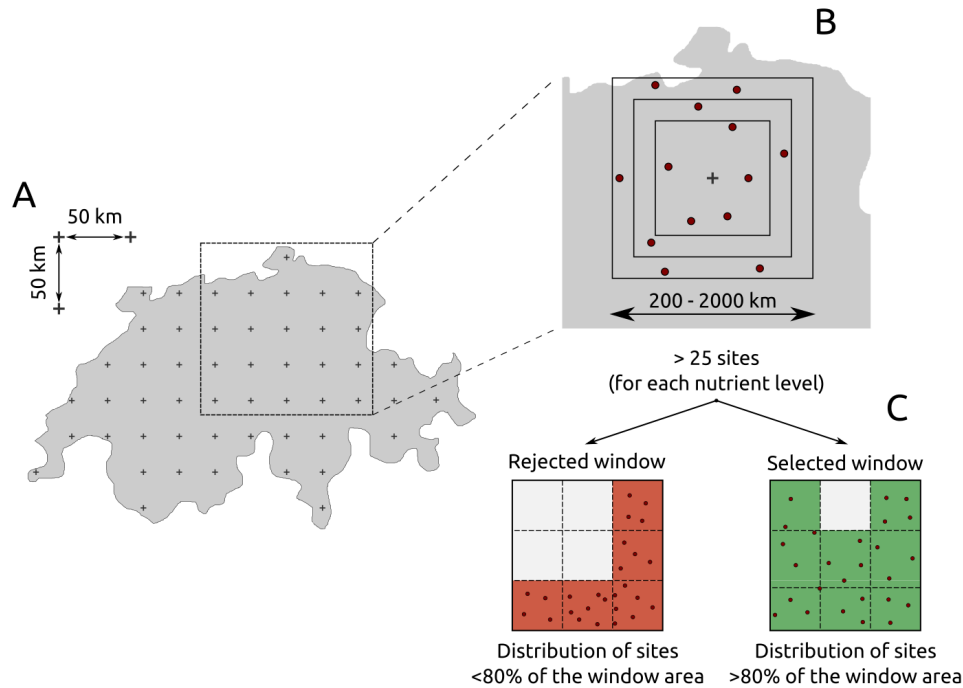


FIGURE 3.2 – Schematic representation of the landscape windows approach. A – The centers of the windows were aligned every 50 km along latitudinal transects in the study area. B – We selected windows that contained at least 25 sites for each nutrient level. C – For each nutrient level, sites must be distributed over at least 80 % of the window's area.

Because total  $\beta$ -diversity can be influenced by sampling effort, we followed the Bennett and Gilbert's recommendation (2016) and randomly sampled 25 sites per window for each impact category to measure  $\beta$ -diversity. We conducted a sensitivity analysis varying sample size from two sites to 250 sites to validate the choice of 25 sites (Appendix S4). We resampled the data 999 times for each window and calculated the corresponding mean  $\beta$ -diversity. The maximum value of the total  $\beta$ -diversity given by the `beta.div.comp` function is 0.5 when sites contain no shared species, and the minimum is 0 when all sites share the same species (Legendre and De Cáceres, 2013).

### Relative influence of local community assembly processes vs. $\gamma$ -diversity

We estimated the contributions of the local community assembly processes vs  $\gamma$ -diversity on  $\beta$ -diversity, across spatial scales, with a null model (Chase and Myers, 2011; Kraft et al., 2011). This procedure randomizes individuals among sites, while preserving  $\gamma$ -

diversity, the total number of individuals per site and the species abundance distributions within each window. After randomizing individuals, we built new presence/absence tables to calculate  $\beta$ -diversity as explained above. We modified this method to be compatible with the Baselga-family decomposition of the Sørensen dissimilarity coefficient for species presence-absence data by transforming each randomized null matrix into a total  $\beta$ -diversity matrix (999 matrices total). Then, we calculated  $\beta$ -deviation ( $\beta_{dev}$ ) with the following equation

$$\beta_{dev} = \frac{\beta_{obs} - \beta_{null}}{sd_{null}}$$

where  $\beta_{obs}$  is the  $\beta$ -diversity of the original sample,  $\beta_{null}$ , the mean  $\beta$ -diversity of the 999 random samples (null model), and  $sd_{null}$ , the standard deviation of  $\beta$ -diversity of the 999 random samples.

Positive values indicate that observed  $\beta$ -diversity is greater than expected if  $\beta$ -diversity were driven solely by  $\gamma$ -diversity, thus highlighting the importance of local community assembly processes.

## Characterization of local community assembly processes

To estimate the relative response of  $\beta$ -diversity to environmental and spatial factors, we tested the correlations between community, environmental and geographical distance matrices with partial Mantel tests in each window for each resampling (999 times) and calculated the mean value for each window (p-values calculated with 999 permutations). Partial Mantel tests have been criticized and we are conscious that they have some limitations since they may in some cases only weakly capture the true effect of space on communities. However, despite recent and intense debates (Laliberté, 2008; Legendre et al., 2005, 2008; Péliissier et al., 2008; Tuomisto and Ruokolainen, 2006), on the relative merits of alternative methods (variance partitioning with PCNM, polynomial RDA or MEM as spatial predictors), Gilbert and Bennett (2010) highlighted that all these spatial methods failed to correctly model the relative influence of environmental and spatial (*e.g.* dispersal effects) drivers on community composition. The partition of the variance is also very sensitive to the number of sites used in calculation, which is a very problematic point in our landscape window frame. Therefore, we considered partial Mantel test as the most appropriate method to analyze our data.

To determine if communities were significantly structured by pure environmental or spatial factors at each scale, we used  $\chi$ -square tests. The partial Mantel tests tracked significant correlations between pure spatial processes (spatial distances) or pure niche/habitat processes (environmental distances) and community distances (*i.e.* Sørensen dissimilarities)

after partialling out the effects of the other explanatory matrix. To characterize the environment, we used all the variables listed earlier, nutrients included. Latitude and longitude were used as spatial predictors.

## Statistical analyses

Given that simulation and permutation procedures generated an artificially large sample size, we analyzed the distributions of  $\beta$ -diversity,  $\beta$ -deviation and partial Mantel statistics with the Cliff's delta statistic (1993) which accounts for large sample size (Tecchio et al., 2016). This statistic estimates the probability that a randomly selected value in one group is higher than a randomly selected value in a second group minus the reverse probability, *i.e.*,  $\delta = P(x_1 > x_2) - P(x_1 < x_2)$ . Comparing the degree of overlap between the two groups indicates whether a significant difference is due to an effective ecological process or to sample size alone. Romano et al. (2016) interpreted the magnitude of differences between groups as follows : negligible for  $|d| < 0.147$ , small for  $0.147 < |d| < 0.33$ , medium for  $0.33 < |d| < 0.474$  and large for higher  $|d|$  values. In this study, we considered  $|d| > 0.33$  as a threshold for significance (see Appendix S2). All data analyses were performed with R 3.4.2 (R Core Team, 2017), using the 'vegan' and 'adespatial' packages (Dray et al., 2018; Oksanen et al., 2018).

## Results

### The effect of scale on $\beta$ -diversity values

Sørensen dissimilarities ranged between 0.27 (heavily impacted sites) and 0.37 (unimpacted sites) in France and between 0.30 (heavily impacted sites) and 0.38 (unimpacted sites) in the US. Although the minimum and maximum values were similar between the two countries,  $\beta$ -diversity patterns across spatial scales were different.  $\beta$ -diversity increased with spatial scale regardless of enrichment level (Fig. 3.3), however, the rate of increase differed across nutrient enrichment levels in both France and the US. In France,  $\beta$ -diversity values in unimpacted sites increased significantly with spatial scale, while in slightly and heavily impacted sites,  $\beta$ -diversity increased to about 400 km. In the US,  $\beta$ -diversity increased significantly with scale to about 1500 km across enrichment levels, but tapered off between 1500 km and 2000 km. Only the 800 km scale was common to France and the US, and it is notable that the corresponding  $\beta$ -diversity values were clearly distinct, especially for unimpacted sites.

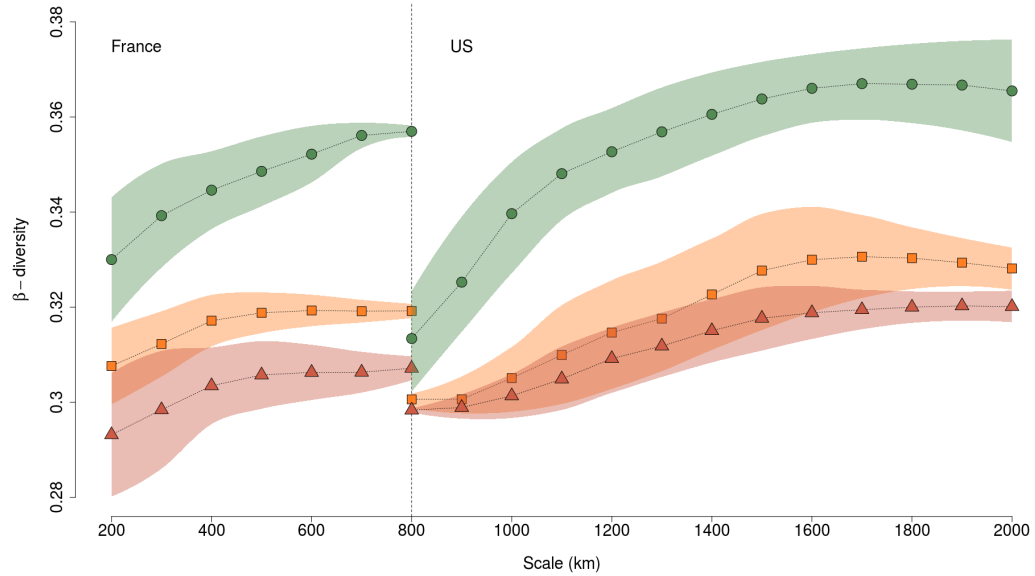


FIGURE 3.3 – Mean  $\beta$ -diversity values (symbols) and SD (bands) measured in unimpacted conditions (round symbols), slightly impacted conditions (square symbols) and heavily impacted conditions (triangle symbols) across spatial scales.

### Impact of nutrient enrichment on $\beta$ -diversity and specialist species

$\beta$ -diversity decreased significantly with nutrient enrichment and even a slight impact had a strong effect on community similarities (Fig. 3.3). Moreover, in France  $\beta$ -diversity tended to plateau at 400 km in impacted sites (slightly and heavily), but continued to increase in unimpacted ones. In the US,  $\beta$ -diversity differences were even more pronounced. For unimpacted sites,  $\beta$ -diversity increased steeply between 800 km and 1200 km, while at the same scales  $\beta$ -diversity in impacted sites increased only slightly. As in France,  $\beta$ -diversity in the US significantly declined with the level of impact, except at 1100 km and 1200 km where there was no difference between slightly and heavily impacted sites.

The proportion of specialist species declined among heavily impacted sites, given the significant increase of  $\theta$ -values (Fridley coefficient) with nutrient enrichment level in both France and the US (Fig. 3.4). Although we observed that for the French data, there was a significant difference between unimpacted sites and slightly impacted sites, such a difference was absent for the US data.

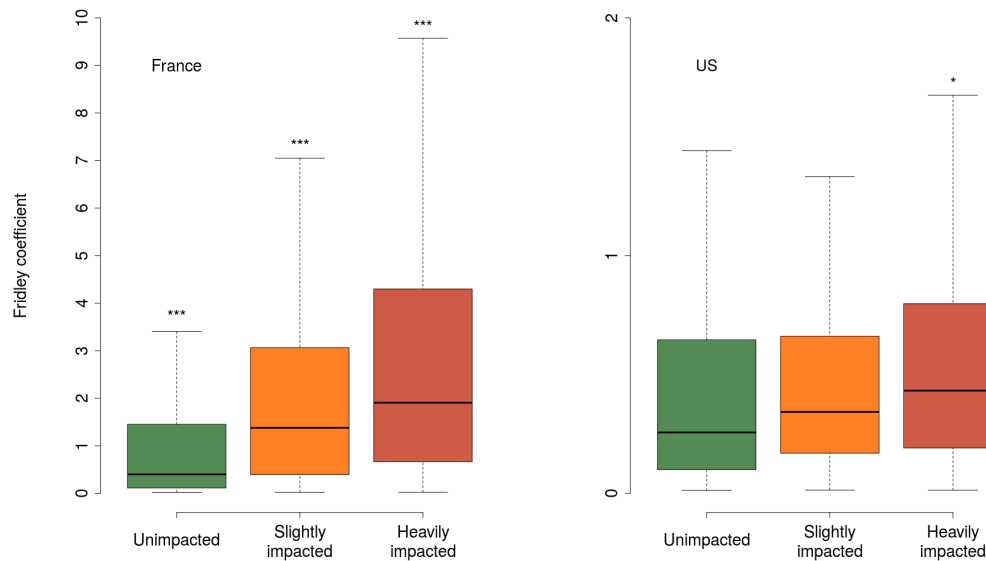


FIGURE 3.4 – Boxplot of the mean  $\theta$ -values (Fridley coefficient) measured in unimpacted, slightly impacted and heavily impacted conditions. Differences between nutrient levels were tested with Kruskal-Wallis tests followed-up by Nemenyi post-hoc tests. Stars represent significant differences between nutrient levels ( $p < 0.001$  : ‘\*\*\*’;  $p < 0.01$  : ‘\*\*’; and  $p < 0.05$  : ‘\*’).

### Relative contributions of local community assembly processes vs. $\gamma$ -diversity on $\beta$ -deviations

The sign and magnitude of  $\beta$ -deviation indicated that  $\beta$ -diversity among diatom communities was driven primarily by local community assembly processes rather than by  $\gamma$ -diversity, regardless of trophic level and spatial scale (Fig. 3.5). Similar to  $\beta$ -diversity,  $\beta$ -deviation varied with scale. In France,  $\beta$ -deviation in unimpacted sites increased significantly from local to larger scales (500 km). In slightly impacted sites, no difference in  $\beta$ -deviation was observed across scales, and in heavily impacted sites, only the first two scales exhibited lower  $\beta$ -deviation values than the larger scales (Fig. 3.5). In the US, across enrichment levels,  $\beta$ -deviation values increased significantly from local scale to 1500 km (impacted sites) or 1600 km (unimpacted sites).

Nutrient enrichment had a strong negative effect on  $\beta$ -deviation, especially at large spatial scales (Fig. 3.5). In France,  $\beta$ -deviation significantly decreased from unimpacted to heavily impacted sites. In the US,  $\beta$ -deviation in unimpacted sites was always significantly higher than in impacted sites (both slightly and heavily). However, except for the 800 km scale, there were no differences between  $\beta$ -deviation values in slightly impacted and heavily impacted sites.

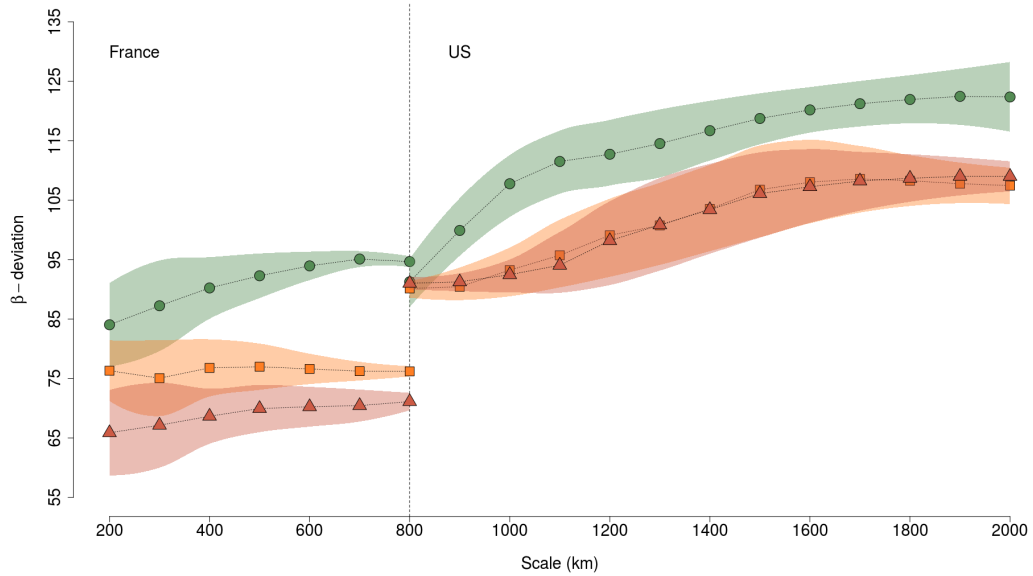


FIGURE 3.5 – Mean  $\beta$ -deviation values (symbols) and SD (bands) measured in unimpacted conditions (round symbols), slightly impacted conditions (square symbols) and heavily impacted conditions (triangle symbols) across spatial scales.

### Characterization of the local community assembly processes

Partial Mantel test results diverged according to impact level (Fig. 3.6). However, there were only very few significant differences between spatial scales and we have not considered them in further discussion (see Appendix S2). In unimpacted sites, there was a significant correlation between diatom  $\beta$ -diversity and pure environmental distance. The pure spatial influence was significant in France at 200 km and 300 km scales, and in the US between 800 km and 1500 km. In contrast, in heavily impacted sites, only pure spatial effect was significant, at small and intermediate scales (200 km to 500 km) in France, and between 1000 km and 2000 km in the US. Pure environmental effect, on the other hand, was non-significant in both France and the US.

Results for slightly impacted sites differed between France and the US. In France, we observed significant correlations between diatom  $\beta$ -diversity and pure environmental distances between 300 km and 700 km. Notably, the pure environmental effect in France was significantly greater in unimpacted sites than in slightly impacted sites regardless of the scale. A pure spatial influence was significant at the 200 km scale only. In the US, pure environmental effect was not significant in slightly impacted sites except at the 2000 km scale, while pure spatial influence was significant across all scales.

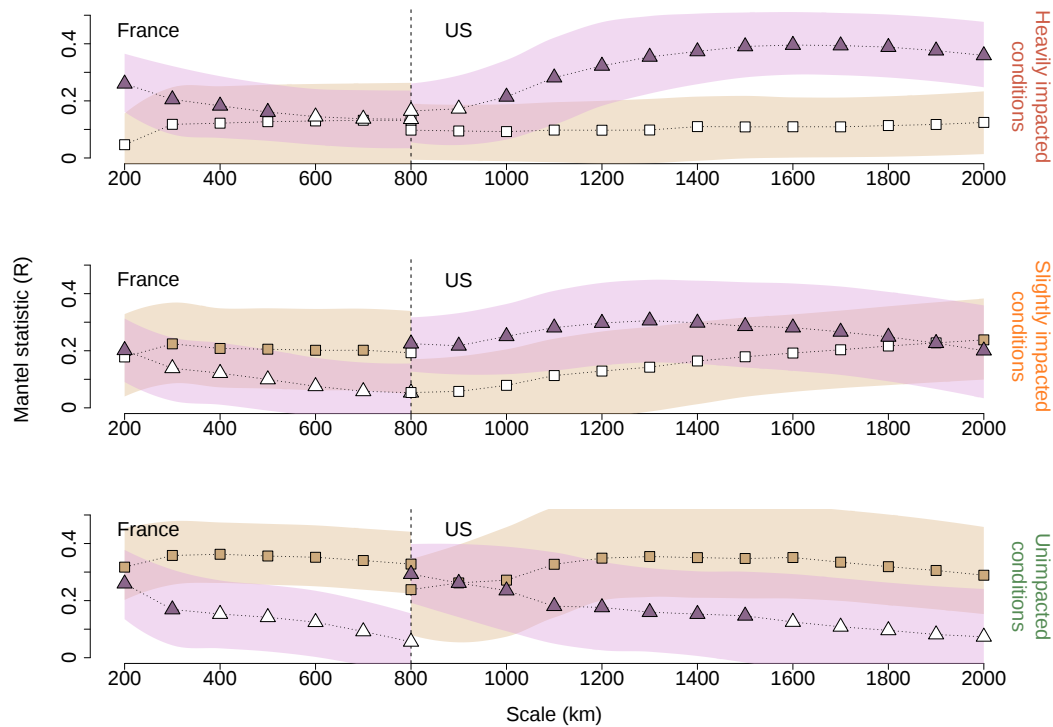


FIGURE 3.6 – Mean partial Mantel coefficients (symbols) and SD (bands) of pure environmental (square symbols) and pure spatial effect (triangle symbols) on  $\beta$ -diversity along a spatial gradient. Filled symbols represent significant relationships (Cliff's  $|d| > 0.33$ ) and open symbols, non-significant ones (Cliff's  $|d| \leq 0.33$ ).

## Discussion

This work provided, for the first time to our knowledge, information on diatom  $\beta$ -diversity patterns and drivers along a nutrient enrichment gradient and across multiple spatial scales. Using an intercontinental dataset and an innovative landscape windows technique, we made the following discoveries. First, we confirmed that  $\beta$ -diversity increases with spatial scale, but nutrient enrichment clearly modulated this increase. Second, in both continents, local assembly processes had a stronger contribution to  $\beta$ -diversity than  $\gamma$ -diversity regardless of scale and nutrient enrichment level. However, nutrient enrichment played a clear role in determining whether species sorting and/or dispersal dominated assembly processes along the spatial gradient. We discuss the implications of these findings below.

## **$\beta$ -diversity across scales and nutrient levels**

We confirmed our first prediction that  $\beta$ -diversity of diatom communities increases with spatial scale, which is generally consistent with prior literature (Gabriel et al., 2006; Martiny et al., 2011). We also revealed that regardless of spatial scale,  $\beta$ -diversity was dependent on nutrient level and that greater enrichment more strongly depressed  $\beta$ -diversity. Anthropogenic and natural stress gradients differ in type and variability, which ultimately drive richness and spatial distribution patterns of sensitive and tolerant species across scales (Clavel et al., 2011). Acute and persistent anthropogenic stress, such as eutrophication, may offer insufficient time for adaptation by specialists and potentially eliminate them from the species pool (Gutiérrez-Cánovas et al., 2013; Kaspari et al., 2010). The loss of specialists benefits subsequent colonization by more physiologically plastic generalists, which are normally poor competitors for resources and niche space compared to specialists (Bolnick et al., 2007; Olden and Rooney, 2006). Our data thus showed that the strength of spatial dependence of  $\beta$ -diversity was likely regulated by how strongly anthropogenic gradients shifted the species pool towards greater prevalence of species with generalist versus specialist traits. Analyses of the Fridley coefficient,  $\theta$ , indeed indicated that these patterns were attributable to the fact that among impacted streams, relatively fewer new species (particularly those with sensitive/specialist traits) were added to the pool compared to unimpacted streams. This result, observed with both continental datasets, supports widely reported observations that agricultural intensification and impact, especially through nutrient enrichment, was a major driver of  $\beta$ -diversity decline, and thus biotic homogenization (Blüthgen et al., 2016; Ekroos et al., 2010; Karp et al., 2012; Vallejós et al., 2016)

However, these findings run counter to studies that reported a positive correlation of  $\beta$ -diversity with productivity (*i.e.* nutrients), but in stochastically assembled communities (Astorga et al., 2014; Chase, 2010). Nutrient levels in our study systems were substantially high (*e.g.* max  $\text{NO}_3 \approx 50 \text{ mg L}^{-1}$ ), and might have exceeded the physiological thresholds of oligotrophic and mesotrophic species, thus resulting in diminished species pool and low  $\beta$ -diversity. Further depression of  $\beta$ -diversity in impacted sites may stem from a reduced influence of stochastic processes, reported for productive systems, which are governed primarily by deterministic species interactions. Notably, our diatom communities were mainly structured by the local environment, and therefore by deterministic processes, as demonstrated below.

## **Role of local assembly in driving $\beta$ -diversity vs $\gamma$ -diversity**

The role of  $\gamma$ -diversity compared to local assembly in driving  $\beta$ -diversity generally increased with impact, given the lower  $\beta$ -diversity values at higher nutrients levels in both datasets. Species pools tended to be more impoverished with increased impact, and domina-

ted by generalists, thus smaller  $\beta$ -deviation at higher impact was likely indicative of increased stochastic sampling of generalists from the regional species pool (Karp et al., 2012; Kraft et al., 2011). Therefore, under stressful conditions,  $\beta$ -diversity was more strongly constrained by regional processes, affecting the distribution of specialists and the size of the regional species pool. A similar finding was reported for diatoms in an acid-impacted region, where acid stress increased the importance of the regional species pool for  $\beta$ -diversity due to the elimination of acid-sensitive species (Pound et al., 2018). Further, we generally observed weak scale dependence of  $\beta$ -deviation in heavily impacted sites. This suggested that the relative roles of  $\gamma$ -diversity and local assembly processes may be decoupled from spatial scale if anthropogenic disturbance is high. Our results disagree to a certain extent with prior research, reporting that the relative importance of community assembly processes in wood plant communities is strongly scale dependent along elevation gradients (Tello et al., 2015). It is thus possible that the scale-dependence of the mechanisms underlying  $\beta$ -diversity may vary along natural (e.g. elevation) vs. anthropogenic gradients (here, eutrophication).

### **Relative influence of species sorting vs dispersal processes**

With heavy nutrient enrichment, the correlation between community dissimilarity and pure environmental distance virtually disappeared whereas the correlation between community dissimilarity and spatial distance increased. This suggested that the influence of species sorting was weakened with nutrient enrichment and dispersal processes became a more important driver of community composition. Mass effects and dispersal limitation are the two principal dispersal processes that can generate spatial structure of community dissimilarity, with mass effects treated as a community homogenizer (Mouquet and Loreau, 2003; Shmida and Wilson, 1985) and dispersal limitation, as a community differentiator (Heino et al., 2015; Hubbell, 2001; Jamoneau et al., 2018). However, it is difficult to disentangle these two processes with Partial Mantel tests because they both can be inferred from correlations between spatial variables and community composition. It is also possible that the spatial effects are due to unmeasured environmental factors, but this may be less of a problem in our study, given that we measured dominant environmental gradients.

The results of the partial Mantel tests (with mainly significant environmental correlations) indicated that species sorting was the dominant mechanism of local assembly in unimpacted sites. In the US, the strength of species sorting increased with spatial scale. Ozkan et al. (2013), working with bird communities, too documented stronger species sorting compared to dispersal processes across spatial scales. Interestingly, local assembly and species sorting increased only weakly with spatial scale in the French data. A likely reason for this difference can be the biogeographical context of the two datasets, with longer latitudinal and longitudinal gradients (*i.e.* study extent) and greater total richness in the US dataset.

In sum, the increase of  $\beta$ -diversity with spatial scale was apparently stronger when species sorting processes dominated but weaker when dispersal processes prevailed. Severity of anthropogenic impact was also important in determining whether species sorting (unimpacted conditions) or dispersal processes (heavily impacted) dominated across spatial scales. However, under slightly enriched conditions, spatial correlations were significant only among US sites regardless of spatial scale whereas environmental correlations were significant for French sites. Thus, the severity of nutrient enrichment, along with biogeographical context, may influence the roles of environmental and dispersal processes in structuring community dissimilarity and ultimately how fast  $\beta$ -diversity increases with scale.

## References

- Astorga, A., Death, R., Death, F., Paavola, R., Chakraborty, M., Muotka, T., 2014. Habitat heterogeneity drives the geographical distribution of beta diversity : the case of New Zealand stream invertebrates. *Ecology and Evolution* 4, 2693–2702.
- Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 19, 134–143.
- Bennett, J., Gilbert, B., 2016. Contrasting beta diversity among regions : how do classical and multivariate approaches compare ? *Global Ecology and Biogeography* 25, 368–377.
- Blüthgen, N., Simons, N.K., Jung, K., Prati, D., Renner, S.C., Boch, S., Fischer, M., Hölzel, N., Klaus, V.H., Kleinebecker, T., Tschapka, M., Weisser, W.W., Gossner, M.M., 2016. Land use imperils plant and animal community stability through changes in asynchrony rather than diversity. *Nature Communications* 7, 10697.
- Bolnick, D.I., Svanbäck, R., Araújo, M.S., Persson, L., 2007. Comparative support for the niche variation hypothesis that more generalized populations also are more heterogeneous. *Proceedings of the National Academy of Sciences* 104, 10075–10079.
- Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26, 211–252.
- Chase, J.M., 2010. Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. *Science* , 1187820.
- Chase, J.M., 2014. Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science* 25, 319–322.
- Chase, J.M., Kraft, N.J.B., Smith, K.G., Vellend, M., Inouye, B.D., 2011. Using null models to disentangle variation in community dissimilarity from variation in alpha-diversity. *Ecosphere* 2, 1–11.
- Chase, J.M., Leibold, M.A., 2003. *Ecological Niches : Linking Classical and Contemporary Approaches*. University of Chicago Press.

- Chase, J.M., Myers, J.A., 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B : Biological Sciences* 366, 2351–2363.
- Clavel, J., Julliard, R., Devictor, V., 2011. Worldwide decline of specialist species : toward a global functional homogenization ? *Frontiers in Ecology and the Environment* 9, 222–228.
- Cliff, N., 1993. Dominance statistics : Ordinal analyses to answer ordinal questions. *Psychological Bulletin* , 494–509.
- Conroy, M.J., Noon, B.R., 1996. Mapping of Species Richness for Conservation of Biological Diversity : Conceptual and Methodological Issues. *Ecological Applications* 6, 763–773.
- Dong, X., Li, B., He, F., Gu, Y., Sun, M., Zhang, H., Tan, L., Xiao, W., Liu, S., Cai, Q., 2016. Flow directionality, mountain barriers and functional traits determine diatom metacomunity structuring of high mountain streams. *Scientific Reports* 6.
- Dray, S., Blanchet, G., Borcard, D., Clappe, S., Jombart, G.G.T., Larocque, G., Legendre, P., Madi, N., Wagner, H.H., 2018. *adespatial* : Moran's eigenvector maps and related methods for the spatial multiscale analysis of ecological communities. R package version 0.1-1.
- Ekroos, J., Heliölä, J., Kuussaari, M., 2010. Homogenization of lepidopteran communities in intensively cultivated agricultural landscapes. *Journal of Applied Ecology* 47, 459–467.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. *Reviews of Geophysics* 45.
- Fridley, J.D., Vandermaast, D.B., Kuppinger, D.M., Manthey, M., Peet, R.K., 2007. Co-occurrence based assessment of habitat generalists and specialists : a new approach for the measurement of niche width. *Journal of Ecology* 95, 707–722.
- Gabriel, D., Roschewitz, I., Tschardtke, T., Thies, C., 2006. Beta Diversity at Different Spatial Scales : Plant Communities in Organic and Conventional Agriculture. *Ecological Applications* 16, 2011–2021.
- Gilbert, B., Bennett, J.R., 2010. Partitioning variation in ecological communities : do the numbers add up ? *Journal of Applied Ecology* 47, 1071–1082.

- Gutiérrez-Cánovas, C., Millán, A., Velasco, J., Vaughan, I.P., Ormerod, S.J., 2013. Contrasting effects of natural and anthropogenic stressors on beta diversity in river organisms. *Global Ecology and Biogeography* 22, 796–805.
- 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. *Freshwater Biology* 60, 845–869.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology* 25, 1965–1978.
- Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- Huston, M.A., 1994. *Biological Diversity : The Coexistence of Species*. Cambridge University Press.
- Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis : A Comparison of Heuristical and Statistical Approaches. *Ecology* 74, 2204–2214.
- Jamoneau, A., Passy, S.I., Soininen, J., Leboucher, T., Tison-Rosebery, J., 2018. Beta diversity of diatom species and ecological guilds : Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology* 63, 62–73.
- Karp, D.S., Rominger, A.J., Zook, J., Ranganathan, J., Ehrlich, P.R., Daily, G.C., 2012. Intensive agriculture erodes beta-diversity at large scales. *Ecology Letters* 15, 963–970.
- Kaspari, M., Stevenson, B.S., Shik, J., Kerekes, J.F., 2010. Scaling community structure : how bacteria, fungi, and ant taxocenes differentiate along a tropical forest floor. *Ecology* 91, 2221–2226.
- Kraft, N.J.B., Comita, L.S., Chase, J.M., Sanders, N.J., Swenson, N.G., Crist, T.O., Stegen, J.C., Vellend, M., Boyle, B., Anderson, M.J., Cornell, H.V., Davies, K.F., Freestone, A.L., Inouye, B.D., Harrison, S.P., Myers, J.A., 2011. Disentangling the Drivers of Beta-Diversity Along Latitudinal and Elevational Gradients. *Science* 333, 1755–1758.
- Lack, D., 1976. *Island biology ; illustrated by the land birds of Jamaica* .
- Laliberté, E., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3232–3237.

- Lecointe, C., Coste, M., Prygiel, J., 1993. Omnidia : software for taxonomy, calculation of diatom indices and inventories management. *Hydrobiologia* 269-270, 509–513.
- Legendre, P., Borcard, D., Peres-Neto, P.R., 2005. Analyzing Beta Diversity : Partitioning the Spatial Variation of Community Composition Data. *Ecological Monographs* 75, 435–450.
- Legendre, P., Borcard, D., Peres-Neto, P.R., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3238–3244.
- Legendre, P., De Cáceres, M., 2013. Beta diversity as the variance of community data : dissimilarity coefficients and partitioning. *Ecology Letters* 16, 951–963.
- Legendre, P., Legendre, L.F.J., 2012. *Numerical Ecology*. Elsevier.
- 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., Gonzalez, A., 2004. The metacommunity concept : a framework for multi-scale community ecology. *Ecology Letters* 7, 601–613.
- Mac Nally, R., Fleishman, E., Bulluck, L.P., Betrus, C.J., 2004. Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography* 31, 917–929.
- Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., Horner-Devine, M.C., 2011. Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences* 108, 7850–7854.
- McKinney, M.L., Lockwood, J.L., 1999. Biotic homogenization : a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* 14, 450–453.
- Mouquet, N., Loreau, M., 2003. Community Patterns in Source Sink Metacommunities. *The American Naturalist* 162, 544–557.
- Myers, J.A., Chase, J.M., Jiménez, I., Jørgensen, P.M., Araujo-Murakami, A., Paniagua-Zambrana, N., Seidel, R., 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters* 16, 151–157.
- Noss, R.F., 1983. A Regional Landscape Approach to Maintain Diversity. *BioScience* 33, 700–706.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. *vegan : Community ecology package*. R package version , 117–118.

- Olden, J., Poff, N., 2003. Toward a Mechanistic Understanding and Prediction of Biotic Homogenization. *The American Naturalist* 162, 442–460.
- Olden, J.D., Poff, N.L., 2004. Ecological Processes Driving Biotic Homogenization : Testing a Mechanistic Model Using Fish Faunas. *Ecology* 85, 1867–1875.
- Olden, J.D., Rooney, T.P., 2006. On defining and quantifying biotic homogenization. *Global Ecology and Biogeography* 15, 113–120.
- Ozkan, K., Svenning, J.C., Jeppesen, E., Webb, T., 2013. Environmental species sorting dominates forest bird community assembly across scales. *Journal of Animal Ecology* 82, 266–274.
- Pélissier, R., Couteron, P., Dray, S., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3227–3232.
- Pound, K., Lawrence, G.B., Passy, S.I., 2018. Beta diversity response to stress severity and heterogeneity in sensitive versus tolerant stream diatoms. *Diversity and Distributions* in press.
- R Core Team, 2017. R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.
- Romano, J., Kromrey, J.D., Coraggio, J., Skowronek, J., 2016. Appropriate statistics for ordinal level data : Should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys. *Annual Meeting of the Florida Association of Institutional Research* , 1–33.
- Shmida, A., Wilson, M.V., 1985. Biological Determinants of Species Diversity. *Journal of Biogeography* 12, 1–20.
- Socolar, J.B., Gilroy, J.J., Kunin, W.E., Edwards, D.P., 2016. How Should Beta-Diversity Inform Biodiversity Conservation ? *Trends in Ecology & Evolution* 31, 67–80.
- Soininen, J., 2007. Environmental and Spatial Control of Freshwater Diatoms - a Review. *Diatom Research* 22, 473–490.
- Soininen, J., Jamoneau, A., Rosebery, J., Passy, S.I., 2016. Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography* 25, 940–950.
- Tecchio, S., Chaalali, A., Raoux, A., Tous Rius, A., Lequesne, J., Girardin, V., Lassalle, G., Cachera, M., Riou, P., Lobry, J., Dauvin, J.C., Niquil, N., 2016. Evaluating ecosystem-

- level anthropogenic impacts in a stressed transitional environment : The case of the Seine estuary. *Ecological Indicators* 61, 833–845.
- Tello, J.S., Myers, J.A., Macía, M.J., Fuentes, A.F., Cayola, L., Arellano, G., Loza, M.I., Torrez, V., Cornejo, M., Miranda, T.B., Jørgensen, P.M., 2015. Elevational gradients in beta-diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLOS ONE* 10, e0121458.
- Tonkin, J.D., Sundermann, A., Jähnig, S.C., Haase, P., 2015. Environmental Controls on River Assemblages at the Regional Scale : An Application of the Elements of Metacomunity Structure Framework. *PLOS ONE* 10, e0135450.
- Tuomisto, H., 2010a. A diversity of beta diversities : straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33, 2–22.
- Tuomisto, H., 2010b. A diversity of beta diversities : straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33, 23–45.
- Tuomisto, H., Ruokolainen, K., 2006. Analyzing or Explaining Beta Diversity ? Understanding the Targets of Different Methods of Analysis. *Ecology* 87, 2697–2708.
- Underwood, A.J., Chapman, M.G., 1996. Scales of spatial patterns of distribution of intertidal invertebrates. *Oecologia* 107, 212–224.
- Vallejos, M.A.V., Padial, A.A., Vitule, J.R.S., 2016. Human-Induced Landscape Changes Homogenize Atlantic Forest Bird Assemblages through Nested Species Loss. *PLOS ONE* 11, e0147058.
- Vanormelingen, P., Verleyen, E., Vyverman, W., 2008. The diversity and distribution of diatoms : from cosmopolitanism to narrow endemism. *Biodiversity and Conservation* 17, 393–405.
- Viana, D.S., Figuerola, J., Schwenk, K., Manca, M., Hobæk, A., Mjelde, M., Preston, C.D., Gornall, R.J., Croft, J.M., King, R.A., Green, A.J., Santamaría, L., 2016. Assembly mechanisms determining high species turnover in aquatic communities over regional and continental scales. *Ecography* 39, 281–288.
- Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. *Nature* 467, 555.
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Vijver, B.V., Jones, V., Flower, R., Roberts, D., Chepurinov, V.A., Kilroy,

- C., Vanormelingen, P., Wever, A.D., 2007. Historical Processes Constrain Patterns in Global Diatom Diversity. *Ecology* 88, 1924–1931.
- Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30, 279–338.
- Whittaker, R.H., 1972. Evolution and Measurement of Species Diversity. *Taxon* 21, 213–251.
- Willis, K.J., Whittaker, R.J., 2002. Species Diversity–Scale Matters. *Science* 295, 1245–1248.
- Wilson, M.V., Shmida, A., 1984. Measuring Beta Diversity with Presence-Absence Data. *Journal of Ecology* 72, 1055–1064.
- Zelený, D., 2009. Cooccurrence based assessment of species habitat specialization is affected by the size of species pool : reply to. *Journal of Ecology* 97, 10–17.

## **Acknowledgements**

Support from the National Science Foundation (grant NSF DEB-1745348 to SP) is gratefully acknowledged. We also thank AFB (French Agency for Biodiversity) and all French Water Agencies for data contribution and financial support. This document has also been carried out with financial support from the French National Research Agency (ANR) in the frame of the Investments for the future Programme, within the Cluster of Excellence COTE (ANR-10-LABX-45). We also thank S. Fattorini, and the three reviewers, for helpful comments which improved the clarity of this work.

## **Data Accessibility Statement**

All environmental and diatom data for the United States are available here : <https://water.usgs.gov/nawqa/>. All environmental and diatom data for France are the property of the French Water Agencies.

## **Biosketch**

**Thibault Leboucher** is a PhD candidate at the French research institute Irstea (Aquatic Ecosystems and Global Changes Research Unit). His research interests relate

to large-scale community ecology, especially freshwater diatom communities.

**William Budnick** is a PhD candidate at the University of Texas – Arlington. His research interests include macroecology and applications thereof for the conservation of aquatic biodiversity, particularly fish and crayfish assemblages.

**Appendix S1 : Number of windows meeting our criteria (minimum 25 sites per category distributed over at least 80% of the window's area) at each scale**

|               | Spatial extent<br>(km) | Area (10 <sup>5</sup> km <sup>2</sup> ) | Number of<br>windows |
|---------------|------------------------|-----------------------------------------|----------------------|
| France        | 200                    | 0.4                                     | 47                   |
|               | 300                    | 0.9                                     | 237                  |
|               | 400                    | 1.6                                     | 308                  |
|               | 500                    | 2.5                                     | 278                  |
|               | 600                    | 3.6                                     | 224                  |
|               | 700                    | 4.9                                     | 144                  |
|               | 800                    | 6.4                                     | 55                   |
| United States | 800                    | 6.4                                     | 47                   |
|               | 900                    | 8.1                                     | 83                   |
|               | 1000                   | 10                                      | 158                  |
|               | 1100                   | 12.1                                    | 300                  |
|               | 1200                   | 14.4                                    | 461                  |
|               | 1300                   | 16.9                                    | 634                  |
|               | 1400                   | 19.6                                    | 718                  |
|               | 1400                   | 19.6                                    | 718                  |
|               | 1500                   | 22.5                                    | 856                  |
|               | 1600                   | 25.6                                    | 1054                 |
|               | 1700                   | 28.9                                    | 1216                 |
|               | 1800                   | 32.4                                    | 1319                 |
|               | 1900                   | 36.1                                    | 1351                 |
|               | 2000                   | 40                                      | 1364                 |

## Appendix S2 : Cliff's delta statistic

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.361  | 0.654  | 0.825  | 0.934  | 0.997  | 1      |
| 300 km |        | 0.323  | 0.541  | 0.704  | 0.841  | 0.857  |
| 400 km |        |        | 0.296  | 0.538  | 0.761  | 0.794  |
| 500 km |        |        |        | 0.294  | 0.62   | 0.67   |
| 600 km |        |        |        |        | 0.388  | 0.48   |
| 700 km |        |        |        |        |        | 0.205  |

TABLE 3.2 –  $\beta$ -diversity – France – Unimpacted sites

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.421  | 0.710  | 0.798  | 0.812  | 0.805  | 0.800  |
| 300 km |        | 0.421  | 0.588  | 0.662  | 0.693  | 0.728  |
| 400 km |        |        | 0.197  | 0.286  | 0.316  | 0.358  |
| 500 km |        |        |        | 0.077  | 0.072  | 0.095  |
| 600 km |        |        |        |        | 0.010  | 0.011  |
| 700 km |        |        |        |        |        | 0.021  |

TABLE 3.3 –  $\beta$ -diversity – France – Slightly impacted sites

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.292  | 0.525  | 0.623  | 0.649  | 0.658  | 0.684  |
| 300 km |        | 0.289  | 0.405  | 0.472  | 0.513  | 0.578  |
| 400 km |        |        | 0.15   | 0.192  | 0.198  | 0.254  |
| 500 km |        |        |        | 0.041  | 0.044  | 0.097  |
| 600 km |        |        |        |        | 0.016  | 0.099  |
| 700 km |        |        |        |        |        | 0.102  |

TABLE 3.4 –  $\beta$ -diversity – France – Heavily impacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.725     | 0.894      | 0.954      | 0.972      | 0.984      | 0.993      | 0.998      | 0.998      | 0.998      | 0.995      | 0.992      | 0.988      |
| 900<br>km  |           | 0.620      | 0.830      | 0.897      | 0.942      | 0.972      | 0.990      | 0.992      | 0.991      | 0.982      | 0.975      | 0.959      |
| 1000<br>km |           |            | 0.414      | 0.606      | 0.743      | 0.849      | 0.934      | 0.961      | 0.959      | 0.933      | 0.901      | 0.842      |
| 1100<br>km |           |            |            | 0.281      | 0.501      | 0.671      | 0.828      | 0.900      | 0.910      | 0.880      | 0.835      | 0.746      |
| 1200<br>km |           |            |            |            | 0.252      | 0.449      | 0.64       | 0.749      | 0.778      | 0.757      | 0.723      | 0.638      |
| 1300<br>km |           |            |            |            |            | 0.221      | 0.433      | 0.565      | 0.615      | 0.605      | 0.591      | 0.526      |
| 1400<br>km |           |            |            |            |            |            | 0.215      | 0.360      | 0.416      | 0.416      | 0.415      | 0.363      |
| 1500<br>km |           |            |            |            |            |            |            | 0.161      | 0.230      | 0.241      | 0.252      | 0.212      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.079      | 0.099      | 0.120      | 0.095      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.023      | 0.048      | 0.030      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.022      | 0.005      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.016      |

TABLE 3.5 –  $\beta$ -diversity – United States – Unimpacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.049     | 0.487      | 0.761      | 0.918      | 0.992      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 900<br>km  |           | 0.482      | 0.760      | 0.924      | 0.994      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 1000<br>km |           |            | 0.397      | 0.619      | 0.757      | 0.833      | 0.871      | 0.899      | 0.920      | 0.937      | 0.939      | 0.925      |
| 1100<br>km |           |            |            | 0.258      | 0.454      | 0.617      | 0.716      | 0.791      | 0.844      | 0.878      | 0.880      | 0.851      |
| 1200<br>km |           |            |            |            | 0.217      | 0.439      | 0.590      | 0.686      | 0.743      | 0.771      | 0.769      | 0.741      |
| 1300<br>km |           |            |            |            |            | 0.290      | 0.521      | 0.664      | 0.741      | 0.770      | 0.764      | 0.735      |
| 1400<br>km |           |            |            |            |            |            | 0.285      | 0.467      | 0.561      | 0.589      | 0.573      | 0.534      |
| 1500<br>km |           |            |            |            |            |            |            | 0.202      | 0.316      | 0.352      | 0.331      | 0.279      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.124      | 0.177      | 0.155      | 0.080      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.070      | 0.050      | 0.052      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.025      | 0.143      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.131      |

TABLE 3.6 –  $\beta$ -diversity – United States – Slightly impacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.214     | 0.32       | 0.72       | 0.933      | 0.993      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 900<br>km  |           | 0.156      | 0.632      | 0.891      | 0.975      | 0.997      | 1          | 1          | 1          | 1          | 1          | 1          |
| 1000<br>km |           |            | 0.388      | 0.655      | 0.785      | 0.860      | 0.909      | 0.944      | 0.969      | 0.984      | 0.984      | 0.981      |
| 1100<br>km |           |            |            | 0.348      | 0.582      | 0.747      | 0.861      | 0.926      | 0.966      | 0.984      | 0.980      | 0.975      |
| 1200<br>km |           |            |            |            | 0.262      | 0.476      | 0.642      | 0.739      | 0.803      | 0.837      | 0.835      | 0.826      |
| 1300<br>km |           |            |            |            |            | 0.243      | 0.467      | 0.620      | 0.727      | 0.797      | 0.803      | 0.789      |
| 1400<br>km |           |            |            |            |            |            | 0.239      | 0.411      | 0.538      | 0.621      | 0.631      | 0.613      |
| 1500<br>km |           |            |            |            |            |            |            | 0.180      | 0.315      | 0.412      | 0.427      | 0.408      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.139      | 0.252      | 0.29       | 0.271      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.124      | 0.183      | 0.166      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.083      | 0.067      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.016      |

TABLE 3.7 –  $\beta$ -diversity – United States – Heavily impacted sites

|         | 200 km | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|---------|--------|--------|--------|--------|--------|--------|--------|
| U - sl  | 0.829  | 0.983  | 0.998  | 1      | 1      | 1      | 1      |
| U - hl  | 0.929  | 0.988  | 1      | 1      | 1      | 1      | 1      |
| sl - hl | 0.641  | 0.682  | 0.868  | 0.891  | 0.959  | 1      | 1      |

TABLE 3.8 –  $\beta$ -diversity – France – Comparison of nutrient levels (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl)

|         | 800<br>km | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|---------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| U - sl  | 1         | 1         | 0.954      | 0.987      | 0.953      | 0.964      | 0.960      | 0.964      | 0.983      | 0.991      | 0.990      | 0.998      | 0.998      |
| U - hl  | 1         | 1         | 0.997      | 1          | 0.998      | 1          | 0.999      | 1          | 1          | 1          | 1          | 1          | 1          |
| sl - hl | 0.757     | 0.686     | 0.566      | 0.303      | 0.249      | 0.331      | 0.467      | 0.562      | 0.678      | 0.815      | 0.897      | 0.881      | 0.843      |

TABLE 3.9 –  $\beta$ -diversity – United States – Comparison of nutrient levels (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.214  | 0.536  | 0.734  | 0.857  | 0.911  | 0.897  |
| 300 km |        | 0.274  | 0.453  | 0.567  | 0.618  | 0.602  |
| 400 km |        |        | 0.261  | 0.462  | 0.573  | 0.544  |
| 500 km |        |        |        | 0.249  | 0.436  | 0.344  |
| 600 km |        |        |        |        | 0.266  | 0.140  |
| 700 km |        |        |        |        |        | 0.193  |

TABLE 3.10 –  $\beta$ -deviation – France – Unimpacted sites

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.094  | 0.040  | 0.053  | 0.040  | 0.041  | 0.081  |
| 300 km |        | 0.156  | 0.160  | 0.116  | 0.065  | 0.061  |
| 400 km |        |        | 0.019  | 0.040  | 0.101  | 0.110  |
| 500 km |        |        |        | 0.038  | 0.076  | 0.061  |
| 600 km |        |        |        |        | 0.040  | 0.010  |
| 700 km |        |        |        |        |        | 0.003  |

TABLE 3.11 –  $\beta$ -deviation – France – Slightly impacted sites

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.194  | 0.430  | 0.576  | 0.629  | 0.641  | 0.673  |
| 300 km |        | 0.277  | 0.451  | 0.516  | 0.574  | 0.687  |
| 400 km |        |        | 0.189  | 0.255  | 0.322  | 0.457  |
| 500 km |        |        |        | 0.045  | 0.113  | 0.276  |
| 600 km |        |        |        |        | 0.059  | 0.216  |
| 700 km |        |        |        |        |        | 0.148  |

TABLE 3.12 –  $\beta$ -deviation – France – Heavily impacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.839     | 0.974      | 0.997      | 0.983      | 0.996      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 900<br>km  |           | 0.700      | 0.905      | 0.891      | 0.928      | 0.978      | 0.996      | 0.998      | 0.999      | 0.999      | 0.998      | 0.981      |
| 1000<br>km |           |            | 0.378      | 0.482      | 0.606      | 0.752      | 0.873      | 0.930      | 0.947      | 0.945      | 0.936      | 0.888      |
| 1100<br>km |           |            |            | 0.194      | 0.353      | 0.565      | 0.746      | 0.851      | 0.884      | 0.886      | 0.870      | 0.806      |
| 1200<br>km |           |            |            |            | 0.188      | 0.412      | 0.622      | 0.753      | 0.809      | 0.828      | 0.820      | 0.758      |
| 1300<br>km |           |            |            |            |            | 0.227      | 0.452      | 0.617      | 0.703      | 0.746      | 0.754      | 0.705      |
| 1400<br>km |           |            |            |            |            |            | 0.241      | 0.432      | 0.556      | 0.630      | 0.663      | 0.624      |
| 1500<br>km |           |            |            |            |            |            |            | 0.215      | 0.372      | 0.48       | 0.543      | 0.529      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.198      | 0.341      | 0.441      | 0.447      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.159      | 0.299      | 0.343      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.160      | 0.238      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.115      |

TABLE 3.13 –  $\beta$ -deviation – United States – Unimpacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.311     | 0.682      | 0.820      | 0.952      | 0.997      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 900<br>km  |           | 0.536      | 0.766      | 0.933      | 0.996      | 1          | 1          | 1          | 1          | 1          | 1          | 1          |
| 1000<br>km |           |            | 0.371      | 0.633      | 0.781      | 0.849      | 0.895      | 0.924      | 0.947      | 0.957      | 0.960      | 0.957      |
| 1100<br>km |           |            |            | 0.311      | 0.516      | 0.664      | 0.776      | 0.851      | 0.904      | 0.925      | 0.930      | 0.923      |
| 1200<br>km |           |            |            |            | 0.219      | 0.431      | 0.605      | 0.710      | 0.777      | 0.804      | 0.811      | 0.800      |
| 1300<br>km |           |            |            |            |            | 0.270      | 0.516      | 0.667      | 0.753      | 0.781      | 0.785      | 0.769      |
| 1400<br>km |           |            |            |            |            |            | 0.308      | 0.495      | 0.592      | 0.613      | 0.607      | 0.587      |
| 1500<br>km |           |            |            |            |            |            |            | 0.195      | 0.304      | 0.326      | 0.317      | 0.293      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.13       | 0.163      | 0.150      | 0.122      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.043      | 0.030      | 0.003      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.012      | 0.048      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.033      |

TABLE 3.14 –  $\beta$ -deviation – United States – Slightly impacted sites

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.158     | 0.348      | 0.483      | 0.871      | 0.983      | 0.998      | 1          | 1          | 1          | 1          | 1          | 1          |
| 900<br>km  |           | 0.211      | 0.401      | 0.826      | 0.969      | 0.992      | 0.999      | 1          | 1          | 1          | 1          | 1          |
| 1000<br>km |           |            | 0.221      | 0.670      | 0.844      | 0.903      | 0.940      | 0.963      | 0.980      | 0.990      | 0.996      | 0.998      |
| 1100<br>km |           |            |            | 0.415      | 0.640      | 0.764      | 0.849      | 0.901      | 0.936      | 0.957      | 0.973      | 0.980      |
| 1200<br>km |           |            |            |            | 0.278      | 0.462      | 0.595      | 0.669      | 0.721      | 0.748      | 0.766      | 0.772      |
| 1300<br>km |           |            |            |            |            | 0.231      | 0.413      | 0.515      | 0.589      | 0.626      | 0.651      | 0.660      |
| 1400<br>km |           |            |            |            |            |            | 0.216      | 0.337      | 0.430      | 0.482      | 0.526      | 0.542      |
| 1500<br>km |           |            |            |            |            |            |            | 0.122      | 0.220      | 0.275      | 0.325      | 0.347      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.101      | 0.152      | 0.195      | 0.214      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.051      | 0.092      | 0.113      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.038      | 0.066      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.040      |

TABLE 3.15 –  $\beta$ -deviation – United States – Heavily impacted sites

|         | 200 km | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|---------|--------|--------|--------|--------|--------|--------|--------|
| U - sl  | 0.678  | 0.789  | 0.952  | 0.994  | 1      | 1      | 1      |
| U - hl  | 0.877  | 0.907  | 0.998  | 1      | 1      | 1      | 1      |
| sl - hl | 0.755  | 0.630  | 0.768  | 0.763  | 0.844  | 0.953  | 1      |

TABLE 3.16 –  $\beta$ -deviation – France – Comparison of nutrient levels (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl)

|         | 800<br>km | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|---------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| U - sl  | 0.139     | 0.998     | 0.936      | 0.962      | 0.834      | 0.865      | 0.831      | 0.767      | 0.841      | 0.895      | 0.951      | 0.970      | 0.927      |
| U - hl  | 0.138     | 0.993     | 0.969      | 0.975      | 0.813      | 0.816      | 0.832      | 0.836      | 0.918      | 0.948      | 0.953      | 0.954      | 0.909      |
| sl - hl | 0.359     | 0.127     | 0.252      | 0.283      | 0.154      | 0.068      | 0.011      | 0.008      | 0.025      | 0.056      | 0.105      | 0.185      | 0.253      |

TABLE 3.17 –  $\beta$ -deviation – United States – Comparison of nutrient levels (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl)

|                   | 200 km | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|-------------------|--------|--------|--------|--------|--------|--------|--------|
| Unimpacted        | 0.275  | 0.726  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| Slightly impacted | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| Heavily impacted  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |

TABLE 3.18 – Partial Mantel tests – France – Pure spatial (SPA) vs. Pure environment (ENV) (not calculated = n.c.)

|    | 800 km | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|----|--------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| U  | 0.330  | 0.057  | 0.090   | 0.428   | 0.523   | 0.605   | 0.618   | 0.613   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| sl | n.c.   | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | 0.145   |
| hl | n.c.   | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |

TABLE 3.19 – Partial Mantel tests – United States – Pure spatial (SPA) vs. Pure environment (ENV) (not calculated = n.c. ; Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl)

|         | 200 km | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|---------|--------|--------|--------|--------|--------|--------|--------|
| U - sl  | n.c.   | 0.546  | 0.622  | 0.606  | 0.599  | 0.560  | n.c.   |
| U - hl  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| sl - hl | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |

TABLE 3.20 – Partial Mantel tests – France – Comparison of nutrient levels – Pure environment (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl ; not calculated = n.c.)

|         | 800 km | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|---------|--------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| U - sl  | n.c.   | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | 0.207   |
| U - hl  | n.c.   | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| sl - hl | n.c.   | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |

TABLE 3.21 – Partial Mantel tests – United States – Comparison of nutrient levels – Pure environment (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl ; not calculated = n.c.)

|         | 200 km | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|---------|--------|--------|--------|--------|--------|--------|--------|
| U - sl  | 0.267  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| U - hl  | 0.010  | 0.311  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| sl - hl | 0.295  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |

TABLE 3.22 – Partial Mantel tests – France – Comparison of nutrient levels – Pure spatial (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl ; not calculated = n.c.)

|         | 800 km | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|---------|--------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| U - sl  | 0.367  | 0.216  | 0.019   | 0.322   | 0.411   | 0.511   | 0.508   | 0.473   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| U - hl  | n.c.   | n.c.   | 0.129   | 0.317   | 0.497   | 0.663   | 0.742   | 0.782   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| sl - hl | n.c.   | n.c.   | 0.183   | 0.001   | 0.107   | 0.199   | 0.307   | 0.420   | 0.450   | 0.490   | 0.527   | 0.548   | 0.568   |

TABLE 3.23 – Partial Mantel tests – United States – Comparison of nutrient levels – Pure spatial (Unimpacted sites = U ; slightly impacted sites = sl ; heavily impacted sites = hl ; not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.372  | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 300 km |        | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 400 km |        |        | n.c.   | n.c.   | n.c.   | n.c.   |
| 500 km |        |        |        | n.c.   | n.c.   | n.c.   |
| 600 km |        |        |        |        | n.c.   | n.c.   |
| 700 km |        |        |        |        |        | n.c.   |

TABLE 3.24 – Partial Mantel tests – France – Pure spatial – Unimpacted sites (not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.189  | 0.205  | 0.175  | 0.152  | 0.097  | 0.037  |
| 300 km |        | 0.013  | 0.019  | 0.044  | 0.103  | 0.165  |
| 400 km |        |        | 0.032  | 0.057  | 0.118  | 0.181  |
| 500 km |        |        |        | 0.025  | 0.086  | 0.150  |
| 600 km |        |        |        |        | 0.060  | 0.125  |
| 700 km |        |        |        |        |        | 0.065  |

TABLE 3.25 – Partial Mantel tests – France – Pure environment – Unimpacted sites (not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 300 km |        | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 400 km |        |        | n.c.   | n.c.   | n.c.   | n.c.   |
| 500 km |        |        |        | n.c.   | n.c.   | n.c.   |
| 600 km |        |        |        |        | n.c.   | n.c.   |
| 700 km |        |        |        |        |        | n.c.   |

TABLE 3.26 – Partial Mantel tests – France – Pure spatial – Slightly impacted sites (not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 300 km |        | 0.066  | 0.075  | 0.088  | 0.089  | n.c.   |
| 400 km |        |        | 0.009  | 0.024  | 0.024  | n.c.   |
| 500 km |        |        |        | 0.014  | 0.015  | n.c.   |
| 600 km |        |        |        |        | 0.001  | n.c.   |
| 700 km |        |        |        |        |        | n.c.   |

TABLE 3.27 – Partial Mantel tests – France – Pure environment – Slightly impacted sites (not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | 0.278  | 0.398  | 0.506  | n.c.   | n.c.   | n.c.   |
| 300 km |        | 0.107  | 0.212  | n.c.   | n.c.   | n.c.   |
| 400 km |        |        | 0.108  | n.c.   | n.c.   | n.c.   |
| 500 km |        |        |        | n.c.   | n.c.   | n.c.   |
| 600 km |        |        |        |        | n.c.   | n.c.   |
| 700 km |        |        |        |        |        | n.c.   |

TABLE 3.28 – Partial Mantel tests – France – Pure spatial – Heavily impacted sites (not calculated = n.c.)

|        | 300 km | 400 km | 500 km | 600 km | 700 km | 800 km |
|--------|--------|--------|--------|--------|--------|--------|
| 200 km | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 300 km |        | n.c.   | n.c.   | n.c.   | n.c.   | n.c.   |
| 400 km |        |        | n.c.   | n.c.   | n.c.   | n.c.   |
| 500 km |        |        |        | n.c.   | n.c.   | n.c.   |
| 600 km |        |        |        |        | n.c.   | n.c.   |
| 700 km |        |        |        |        |        | n.c.   |

TABLE 3.29 – Partial Mantel tests – France – Pure environment – Heavily impacted sites (not calculated = n.c.)

|         | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|---------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 800 km  | 0.151  | 0.209   | 0.400   | 0.445   | 0.531   | 0.550   | 0.552   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 900 km  |        | 0.078   | 0.275   | 0.311   | 0.393   | 0.417   | 0.427   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1000 km |        |         | 0.190   | 0.218   | 0.290   | 0.312   | 0.324   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1100 km |        |         |         | 0.022   | 0.087   | 0.111   | 0.131   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1200 km |        |         |         |         | 0.070   | 0.096   | 0.118   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1300 km |        |         |         |         |         | 0.029   | 0.055   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1400 km |        |         |         |         |         |         | 0.027   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1500 km |        |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1600 km |        |         |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    |
| 1700 km |        |         |         |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    |
| 1800 km |        |         |         |         |         |         |         |         |         |         | n.c.    | n.c.    |
| 1900 km |        |         |         |         |         |         |         |         |         |         |         | n.c.    |

TABLE 3.30 – Partial Mantel tests – United States – Pure spatial – Unimpacted sites (not calculated = n.c.)

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.053     | 0.076      | 0.291      | 0.381      | 0.421      | 0.425      | 0.424      | 0.458      | 0.413      | 0.365      | 0.318      | 0.253      |
| 900<br>km  |           | 0.018      | 0.207      | 0.283      | 0.311      | 0.308      | 0.305      | 0.328      | 0.283      | 0.237      | 0.194      | 0.138      |
| 1000<br>km |           |            | 0.163      | 0.227      | 0.245      | 0.236      | 0.229      | 0.242      | 0.197      | 0.152      | 0.113      | 0.065      |
| 1100<br>km |           |            |            | 0.061      | 0.074      | 0.061      | 0.052      | 0.057      | 0.006      | 0.044      | 0.087      | 0.139      |
| 1200<br>km |           |            |            |            | 0.011      | 0.004      | 0.014      | 0.012      | 0.067      | 0.119      | 0.165      | 0.220      |
| 1300<br>km |           |            |            |            |            | 0.017      | 0.029      | 0.029      | 0.087      | 0.141      | 0.190      | 0.249      |
| 1400<br>km |           |            |            |            |            |            | 0.012      | 0.012      | 0.071      | 0.127      | 0.178      | 0.239      |
| 1500<br>km |           |            |            |            |            |            |            | 0.001      | 0.058      | 0.115      | 0.166      | 0.229      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.061      | 0.120      | 0.175      | 0.242      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.060      | 0.115      | 0.184      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.056      | 0.126      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.071      |

TABLE 3.31 – Partial Mantel tests – United States – Pure environment – Unimpacted sites (not calculated = n.c.)

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | 0.043     | 0.103      | 0.235      | 0.307      | 0.346      | 0.321      | 0.267      | 0.246      | 0.182      | 0.102      | 0.004      | 0.108      |
| 900<br>km  |           | 0.140      | 0.267      | 0.336      | 0.374      | 0.349      | 0.296      | 0.275      | 0.211      | 0.133      | 0.035      | 0.078      |
| 1000<br>km |           |            | 0.126      | 0.196      | 0.233      | 0.208      | 0.158      | 0.139      | 0.079      | 0.006      | 0.084      | 0.186      |
| 1100<br>km |           |            |            | 0.070      | 0.107      | 0.081      | 0.035      | 0.018      | 0.039      | 0.108      | 0.191      | 0.284      |
| 1200<br>km |           |            |            |            | 0.037      | 0.010      | 0.034      | 0.050      | 0.104      | 0.171      | 0.250      | 0.337      |
| 1300<br>km |           |            |            |            |            | 0.026      | 0.070      | 0.085      | 0.139      | 0.205      | 0.281      | 0.366      |
| 1400<br>km |           |            |            |            |            |            | 0.044      | 0.059      | 0.113      | 0.180      | 0.257      | 0.343      |
| 1500<br>km |           |            |            |            |            |            |            | 0.016      | 0.070      | 0.136      | 0.214      | 0.301      |
| 1600<br>km |           |            |            |            |            |            |            |            | 0.054      | 0.119      | 0.196      | 0.283      |
| 1700<br>km |           |            |            |            |            |            |            |            |            | 0.065      | 0.143      | 0.232      |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | 0.080      | 0.171      |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | 0.092      |

TABLE 3.32 – Partial Mantel tests – United States – Pure spatial – Slightly impacted sites (not calculated = n.c.)

|         | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|---------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 800 km  | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 900 km  |        | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1000 km |        |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1100 km |        |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1200 km |        |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1300 km |        |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1400 km |        |         |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1500 km |        |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1600 km |        |         |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    | n.c.    |
| 1700 km |        |         |         |         |         |         |         |         |         | n.c.    | n.c.    | n.c.    |
| 1800 km |        |         |         |         |         |         |         |         |         |         | n.c.    | n.c.    |
| 1900 km |        |         |         |         |         |         |         |         |         |         |         | n.c.    |

TABLE 3.33 – Partial Mantel tests – United States – Pure environment – Slightly impacted sites (not calculated = n.c.)

|         | 900 km | 1000 km | 1100 km | 1200 km | 1300 km | 1400 km | 1500 km | 1600 km | 1700 km | 1800 km | 1900 km | 2000 km |
|---------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 800 km  | n.c.   | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 900 km  |        | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    | n.c.    |
| 1000 km |        |         | 0.265   | 0.422   | 0.531   | 0.607   | 0.665   | 0.682   | 0.679   | 0.664   | 0.630   | 0.584   |
| 1100 km |        |         |         | 0.164   | 0.286   | 0.365   | 0.435   | 0.454   | 0.448   | 0.428   | 0.382   | 0.320   |
| 1200 km |        |         |         |         | 0.128   | 0.205   | 0.281   | 0.301   | 0.293   | 0.270   | 0.217   | 0.147   |
| 1300 km |        |         |         |         |         | 0.074   | 0.152   | 0.172   | 0.162   | 0.138   | 0.081   | 0.007   |
| 1400 km |        |         |         |         |         |         | 0.082   | 0.102   | 0.092   | 0.066   | 0.005   | 0.074   |
| 1500 km |        |         |         |         |         |         |         | 0.020   | 0.008   | 0.018   | 0.081   | 0.163   |
| 1600 km |        |         |         |         |         |         |         |         | 0.012   | 0.039   | 0.102   | 0.184   |
| 1700 km |        |         |         |         |         |         |         |         |         | 0.027   | 0.091   | 0.174   |
| 1800 km |        |         |         |         |         |         |         |         |         |         | 0.063   | 0.147   |
| 1900 km |        |         |         |         |         |         |         |         |         |         |         | 0.083   |

TABLE 3.34 – Partial Mantel tests – United States – Pure spatial – Heavily impacted sites (not calculated = n.c.)

|            | 900<br>km | 1000<br>km | 1100<br>km | 1200<br>km | 1300<br>km | 1400<br>km | 1500<br>km | 1600<br>km | 1700<br>km | 1800<br>km | 1900<br>km | 2000<br>km |
|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 800<br>km  | n.c.      | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 900<br>km  |           | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1000<br>km |           |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1100<br>km |           |            |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1200<br>km |           |            |            |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1300<br>km |           |            |            |            |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1400<br>km |           |            |            |            |            |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1500<br>km |           |            |            |            |            |            |            | n.c.       | n.c.       | n.c.       | n.c.       | n.c.       |
| 1600<br>km |           |            |            |            |            |            |            |            | n.c.       | n.c.       | n.c.       | n.c.       |
| 1700<br>km |           |            |            |            |            |            |            |            |            | n.c.       | n.c.       | n.c.       |
| 1800<br>km |           |            |            |            |            |            |            |            |            |            | n.c.       | n.c.       |
| 1900<br>km |           |            |            |            |            |            |            |            |            |            |            | n.c.       |

TABLE 3.35 – Partial Mantel tests – United States – Pure environment – Heavily impacted sites  
(not calculated = n.c.)

## **Appendix S3 : Landscape windows functions**

These functions create landscape windows into an area defined by a shapefile (shape format). There are two different methods : regular and random. Using regular method : latitudinal transects aligned every n kilometers are used to place the center of each windows. Windows are aligned along these transects at n kilometers distances between their centers. Using random method : n windows are placed randomly in the defined area. With default values, to be validated, a window must contain at least 25 sites distributed over at least 80% of the window's area (at least 1 site in 8 cells of the 3 x 3 grid).

### **How to install the landscape windows package ?**

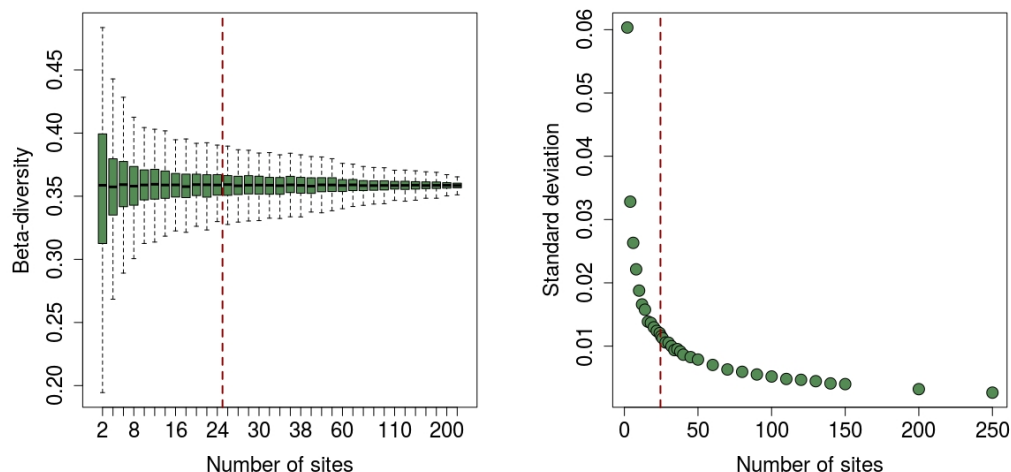
```
devtools::install_github("tleboucher/landscape.windows")
```

## Appendix S4 : Sensitivity analysis

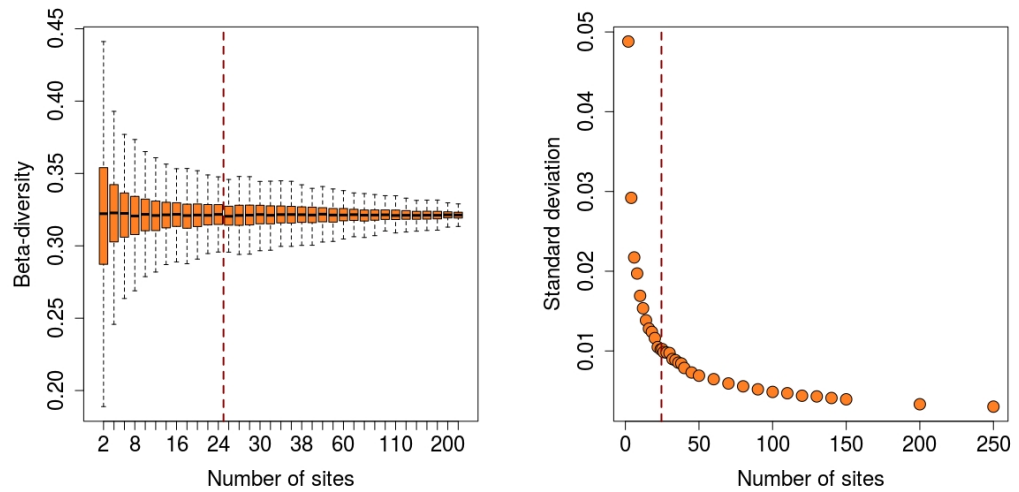
The choice to measure  $\beta$ -diversity on a selection of 25 sites is the result of Bennett and Gilbert's recommendation (2016) who consider that it is necessary to have a minimum of 25 sites to make reliable  $\beta$ -diversity measurements, and of a compromise between the number of sites and the number of landscape windows (the number of validated landscape windows decreases with the increase in the number of sites required to measure  $\beta$ -diversity). In order to validate our choice, we conducted a sensitivity analysis by measuring  $\beta$ -diversity, in France and the US separately for each nutrient level, including an increasing number of sites (from 2 sites to 250 sites). We repeated the measurement 999 times for each number of sites. The graphs below represent the  $\beta$ -diversity values obtained in function of the number of sites (left) and the standard deviation of  $\beta$ -diversity in function of the number of sites (right) in France and in the US, for each nutrient level. The vertical dotted line corresponds to the number of sites chosen (25).

### France

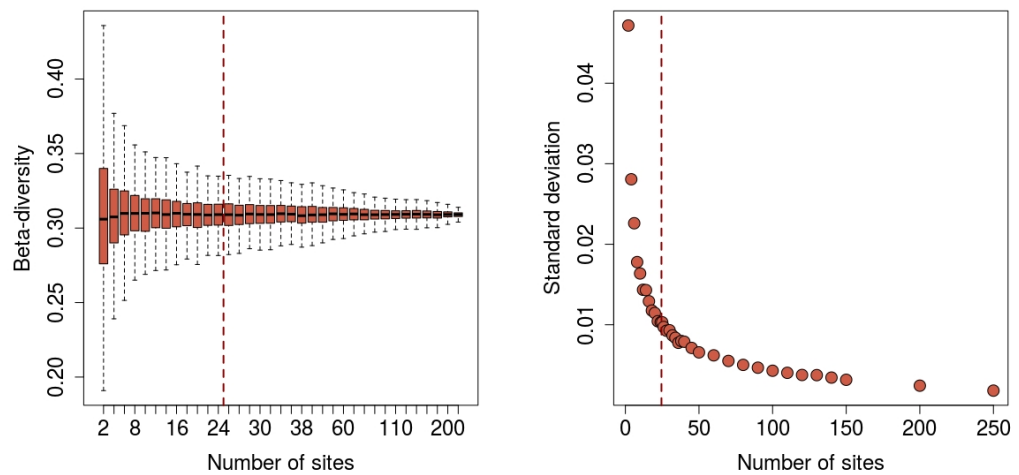
#### Unimpacted sites – Number of sites = 636



### Slightly impacted sites – Number of sites = 1163

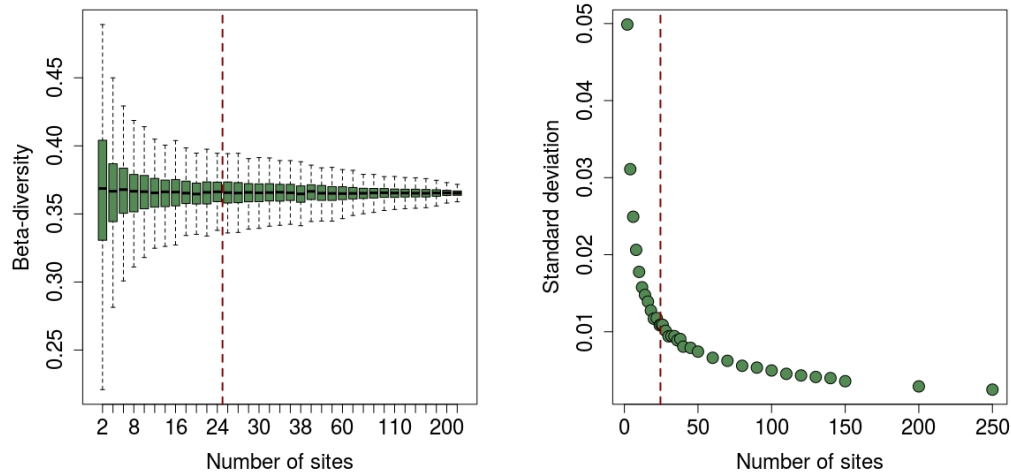


### Heavily impacted sites – Number of sites = 413

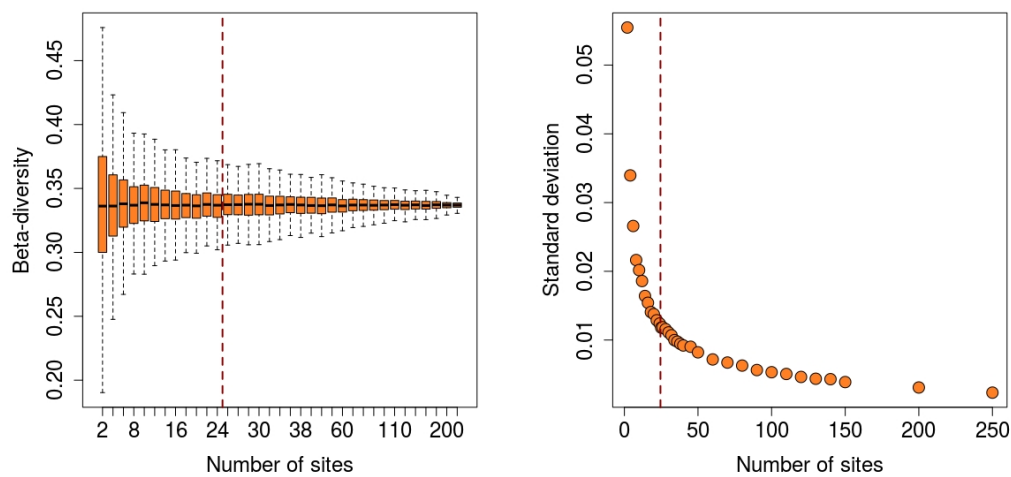


United-States

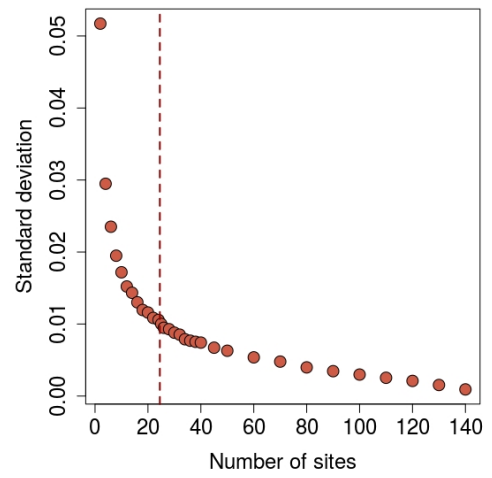
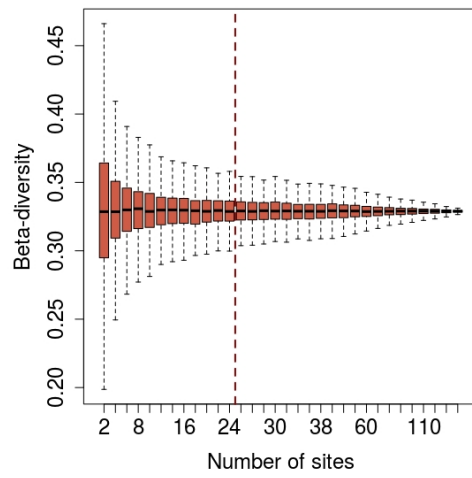
Unimpacted sites – Number of sites = 555



Slightly impacted sites – Number of sites = 358



## Heavily impacted sites – Number of sites = 165







# 4 Développement méthodologique pour la mise en évidence de l'influence de l'effet de masse sur la distribution des espèces

Ce chapitre présente un article en cours de finalisation qui sera soumis à la revue *Journal of Applied Ecology*.

## 4.1 Contexte et objectifs de l'étude

Le chapitre précédent a permis de mettre en évidence qu'une structuration purement spatiale des communautés de diatomées existait mais sans révéler quel processus était véritablement à l'œuvre. La variabilité entre les capacités de dispersion des espèces est à l'origine de la structuration spatiale de toutes communautés. Si l'existence de tels processus est connu des scientifiques depuis longtemps (MacArthur and Wilson, 1967), ceux-ci ont longtemps été sous-estimés. Ce n'est qu'à la fin du XX<sup>ème</sup> siècle que les écologues ont commencé à reconsidérer les processus basés sur les capacités de dispersion des espèces au sein de leurs théories (Hubbell, 2001; Leibold et al., 2004; Palmer et al., 1996). Ainsi la théorie des méta-communautés (Leibold et al., 2004) définit que les communautés locales sont connectées au sein d'une région et sont de fait sujettes aux processus locaux (comme les filtres environnementaux ou les interactions biotiques) mais également aux processus régionaux de dispersion et aux interactions entre ces deux types de processus (Amarasekare and Nisbet, 2001; Chase, 2014; Holyoak et al., 2005).

L'effet de masse et la dispersion limitée sont les deux principaux processus de dispersion agissant sur les communautés de diatomées. Comme défini précédemment, l'effet de masse est un processus asymétrique permettant la présence d'une espèce dans un habitat qui lui est peu favorable (puits) via l'immigration d'individus vivant dans des sites où l'environnement est favorable à l'espèce (sources) (Mouquet and Loreau, 2003; Shmida

and Wilson, 1985). A l'inverse, la dispersion limitée correspond à l'incapacité d'une espèce à atteindre tous les habitats qui lui sont favorables, à cause de capacités de dispersion insuffisantes ou de barrières physiques, qu'elles soient naturelles ou non (Heino et al., 2015). Si la dispersion limitée peut être appréhendée par le biais de démarches expérimentales (Ehrlén and Eriksson, 2000) ou de modélisation (Svenning et al., 2008), l'effet de masse est plus difficile à mettre en évidence en particulier chez les disperseurs passifs que sont les diatomées. Des approches combinant signaux spatiaux et traits signant les capacités de dispersion des espèces ne peuvent ainsi pas être utilisés, du fait de mettre ainsi en lumière l'importance de sa prise en compte lors l'absence totale d'information sur ces seconds. Cela complique grandement notre capacité à démêler l'influence de l'effet de masse et de la dispersion limitée puisque ces deux processus sont classiquement mis en évidence par des signaux spatiaux.

Dans ce chapitre, nous avons cherché à surmonter ce problème en développant une méthode innovante basée sur une combinaison d'analyses : les Asymmetric Eigenvector Maps et les patrons de co-occurrences négatives.

## 4.2 Principaux résultats

Ce travail se base sur l'exploitation des données de quatre grands bassins versants en France, la Garonne, la Seine, la Loire et le Rhône, regroupant respectivement 223, 73, 404 et 254 sites distincts. Les données ont été récoltées entre le 15 juin et le 15 août 2012, afin de limiter l'effet des successions écologiques et l'effet de perturbations extérieures.

La combinaison de méthodes innovantes nous a permis de mettre en évidence trois groupes de profils distincts, dont le troisième se caractérise par de faibles valeurs de co-occurrences négatives, une forte corrélation aux facteurs purement spatiaux ainsi que des valeurs de tolérance moyenne à faible. Ces caractéristiques correspondent à ce qui était attendu d'espèces très sensibles à l'effet de masse.

C'est ce dernier groupe qui nous a permis d'établir une liste d'une cinquantaine d'espèces dont la distribution était influencée par l'effet de masse. Considérant le fait que la présence d'espèces dans des habitats qui leur sont *a priori* peu favorables peut potentiellement masquer l'effet des conditions locales du milieu sur les communautés, et ainsi induire un biais dans l'évaluation de l'état écologique des cours d'eau, cette liste d'espèces devra faire l'attention d'une attention particulière dans les méthodes de bio-indications basées sur les diatomées, en leur donnant un poids moins important dans les calculs des indices, ou en les retirant simplement de ces calculs.

Si l'effet de masse avait déjà été évoqué par différentes études auparavant (voir Bottin

et al., 2016; Goldenberg Vilar et al., 2014; Jamoneau et al., 2018), c'est à notre connaissance la première fois qu'il est clairement identifié sans confusion possible avec la dispersion limitée.

# **Asymmetric eigenvector maps and co-occurrence patterns provide robust evidence for mass effect influence on stream diatoms**

**Thibault Leboucher<sup>1\*</sup>, Juliette Tison-Rosebery<sup>1</sup>, William R. Budnick<sup>2</sup>, Aurélien Jamoneau<sup>1</sup>, Wim Vyverman<sup>3</sup>, Janne Soininen<sup>4</sup>, Sébastien Boutry<sup>1</sup> and Sophia I. Passy<sup>2</sup>**

<sup>1</sup> Aquatic Ecosystems and Global Changes Research Unit, IRSTEA, 50 avenue de Verdun, 33612 Cestas, France

<sup>2</sup> Department of Biology, University of Texas at Arlington, Box 19498, Arlington, TX 76019, USA

<sup>3</sup> Ghent University, Biology Department, Protistology and Aquatic Ecology, Krijgslaan 281 S8, 9000 Gent, Belgium

<sup>4</sup> Department of Geosciences and Geography, University of Helsinki, PO Box 64, Helsinki, FIN-00014, Finland

\* Corresponding author : thibault.leboucher@irstea.fr

## **Abstract**

1. Mass effect, that allows species to reach unfavourable habitats, and dispersal limitation, that prevents species to reach all favourable habitats, are the two major dispersal processes. While dispersal limitation can be detected by experimental or modeling techniques, mass effect is more challenging to detect, which hampers our ability to disentangle the influence of environmental vs. dispersal mechanisms.
2. We developed an innovative method for the detection of species influenced by mass effect. We tested it on stream diatoms from four French watersheds. This method combined (1) an appropriate spatial framework (Asymmetric Eigenvector Map) used in species distribution modeling to measure the relative influence of dispersal vs. niche processes, (2) an analysis of negative co-occurrence patterns to disentangle mass effect from dispersal limitation, and (3) a measurement of niche breadths in order to ensure the distinction between non-spatially structured generalists and species influenced by mass effect.

3. We classified species into three homogeneous groups, the third of which is characterized by low negative co-occurrence values, a high correlation to purely spatial factors and average to low tolerance values. These characteristics correspond to what was expected of species highly sensitive to mass effect.
4. AEM provided an adequate spatial framework taking into account the among-site connectivity and the particularities of river networks. Our approach linked mass effect and negative co-occurrence patterns and allowed us to identify a group of species that we consider to be influenced by mass effect. The presence of such species in unsuitable habitats may in fact mask the effect of local environmental conditions and result in potentially erroneous evaluation of water quality.
5. *Synthesis and applications* : We suggest that the sensitivity of species towards mass effect should represent a new ecological trait to be considered for fundamental and applied issues concerning ecology and water quality assessment. Almost all of the species identified here as influenced by mass effect are contributing to the calculation of different diatom-based indices (*e.g.* BDI or SPI) and then should be treated with caution when assigning ecological status classes to water bodies.

**Keywords** : Asymmetric Eigenvector Map, bio-assessment, community assembly, diatoms, dispersal, mass effect

## Introduction

Variability in dispersal capacities among species leads to spatial structuring of ecological communities. Even if the existence of such process is familiar to scientists for a long time (MacArthur and Wilson, 1967), it has largely been underestimated until the end of the last century when ecologists reconsidered to integrate dispersion-driven processes into their theoretical concepts and field investigations (Hubbell, 2001; Leibold et al., 2004; Palmer et al., 1996). According to the metacommunity concept (Leibold et al., 2004), local communities within a region are connected via dispersal and are, therefore, subject to both local community processes (*i.e.* environmental filtering, species interactions and ecological drift), regional dispersal effects, and their interaction (Amarasekare and Nisbet, 2001; Chase, 2014; Chase et al., 2005; Holyoak et al., 2005). Both environmental gradients or patches (Bell et al., 2006) and dispersal (Jamoneau et al., 2018; Soininen, 2007; Soininen et al., 2007; Tuomisto et al., 2003) thus were shown to explain community composition in both terrestrial and aquatic communities.

Mass effect and dispersal limitation are the two major dispersal processes. Mass effect is an asymmetric process, involving the dispersal of species from locations with higher population densities to locations with lower population densities. Mass effects are implicated in source-sink dynamics, whereby species are supported in unfavourable conditions (sinks) via immigration from source populations inhabiting favorable environments (Mouquet and Loreau, 2003; Shmida and Wilson, 1985). In contrast, dispersal limitation prevents species from reaching suitable habitats and establishing viable populations due to low dispersal rates or natural barriers (Heino et al., 2015). The influence of these two processes is inferred from spatial analyses. However, while dispersal limitation can be detected by experimental (Ehrlén and Eriksson, 2000) or modeling techniques (Svenning et al., 2008), mass effect is more challenging to detect. The importance of mass effect in structuring active disperser communities, *e.g.* benthic macro-invertebrates, could be assessed by a combination of spatial and trait analyses (Brown and Swan, 2010). However, considering the absence of such traits clearly measuring dispersal abilities, this type of approach cannot be used for passive dispersers, which hampers our ability to disentangle the influence of environmental vs. dispersal mechanisms. As both dispersal limitation and mass effect generate spatial structure and are thus difficult to sort out, the identity of each mechanism is currently inferred from the scale of observation. Spatial structure at small scales was taken to represent mass effect, while at large scales, the spatial signal was attributed to dispersal limitation (Bennett et al., 2010; Ng et al., 2009).

Spatial processes that shape metacommunities also depend on the physical structure of the landscape. For example, dispersal may be constrained in dendritic ecological networks, such as rivers where only limited dispersal pathways exist (Brown et al., 2011; Grant et al., 2007; Liu et al., 2013). Therefore, assessing the spatial structure of such networks cannot rely on commonly used spatial models assuming unrestricted overland dispersal (*e.g.* distance-based Moran eigenvector maps). Asymmetric Eigenvector Maps,

on the other hand, can accurately quantify spatial community structure in the presence of a directional dispersal, *e.g.* within river networks (Blanchet et al., 2008a), and could therefore capture a spatial signal due to mass effect. Nevertheless, the detection of a directional signal does not negate the presence of a non-directional one and cannot exclude the influence of dispersal limitation. Moreover, spatial signals may be due to spatially structured but unmeasured environmental variables (Blanchet et al., 2008a; Landeiro et al., 2011). AEM alone are, therefore, not sufficient to detect the influence of mass effect and this spatial framework should be combined with other analyses, as described below.

As mass effect can allow species to persist in unfavorable habitats (*i.e.* far away from their optimum), it tends to increase niche overlap compared to species primarily driven by niche processes. In contrast, dispersal limitation tends to decrease niche overlap by preventing species from reaching all suitable habitats. Niche overlap can be inferred from an analysis of species co-occurrence patterns, which has been broadly used in this context (*e.g.* Bottin et al., 2016; Larsen and Ormerod, 2014). Thus, strong negative associations suggest prevalence of niche processes on species distribution, while weak negative associations indicate a strong influence of mass effect on species distribution.

The AEM spatial framework combined with negative co-occurrence patterns may provide a solid approach for estimating the influence of mass effect on species distribution, given that only mass effect may induce both directional spatial signals and weak negative co-occurrence patterns (Fig. 4.1). Both specialist and generalist species could be subjects of mass effect. Specialists may dominate in suitable environments and evade filtering in unsuitable habitats in the presence of mass effect. Generalists may trade off dominance in favorable conditions for tolerance in unfavorable environments. Thus, these species thrive in unfavorable environments but may persist in favorable habitats, where they are at a competitive disadvantage, if rescued by mass effect. However, generalists could also display weak negative co-occurrence due to broader environmental niches. Therefore, it is important to distinguish whether generalist distributions are driven by environmental as opposed to dispersal mechanisms. Our approach to detect mass effect influence, illustrated in figure 4.2, includes a combination of three distinct analyses : (1) an appropriate spatial framework (AEM) used in species distribution modeling to measure the relative influence of dispersal vs. niche processes, (2) an analysis of negative co-occurrence patterns to disentangle mass effect from dispersal limitation, and (3) a measurement of niche breadths in order to ensure the distinction between non-spatially structured generalists and species influenced by mass effect.

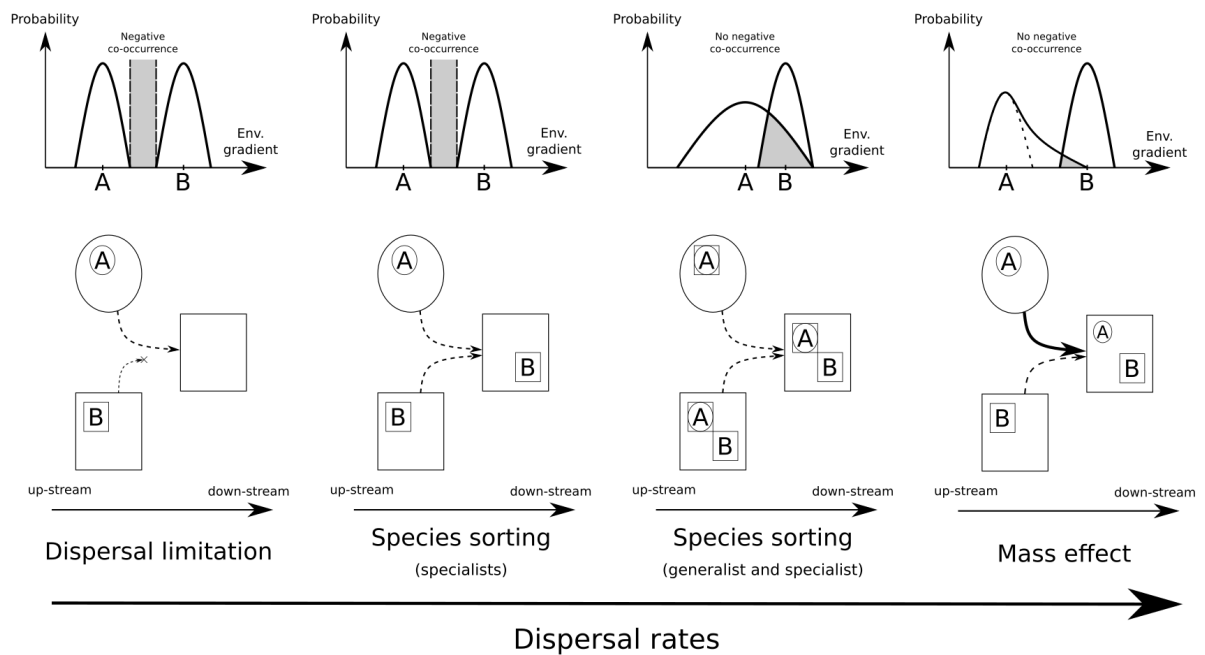


FIGURE 4.1 – Schematic representation of some paradigms of metacommunity theory (adapted from Leibold et al., 2004), given the constraints of a river network. Arrows connect up-stream donor populations of species A and B with potential colonization sites downstream, shown as large ovals and boxes representing two different habitat types. Solid arrows represent higher dispersal than dashed arrows. All arrows are unidirectional, constrained by flow direction (upstream to downstream). The degree to which a species is the competitive dominant in a site is shown by the matching of the smaller box or oval (denoting its habitat preference) with the site symbol. The paradigms illustrated are arranged along a gradient of dispersal and represent dispersal limitation, species sorting with two examples, one with specialist species and one with generalist and specialist species, and mass effect. The niche breadth of the two species associated with each paradigm illustrate the effect of these paradigms on negative co-occurrences patterns.

We tested our approach on microscopic passive dispersers, stream diatoms, which are important contributors to primary production in freshwater ecosystems. Many authors demonstrated that those microalgae were structured by both environmental filtering (Finlay, 2002; Verleyen et al., 2009) and dispersal (Göthe et al., 2013; Jamoneau et al., 2018; Soininen et al., 2016; Vanormelingen et al., 2008; Vyverman et al., 2007). Although mass effect has been evoked as a driver of diatom community patterns (Bottin et al., 2016; Goldenberg Vilar et al., 2014; Jamoneau et al., 2018), this notion has not been conclusively proven due to the lack of a robust technique to distinguish dispersal limitation and mass effect. Convincing evidence is therefore needed since diatoms are widely used in biomonitoring methods around the world. The presence of species in unsuitable habitats

due to mass effect could mask the effect of local environmental conditions resulting in an improper evaluation of water quality. Thus, the major goals of our study were to develop a method for detecting mass effect and determine which species could be influenced by it, and consequently considered with caution in ecological river assessments. The influence of mass effect on particular ecological traits was also examined (e.g. biovolumes and guilds).

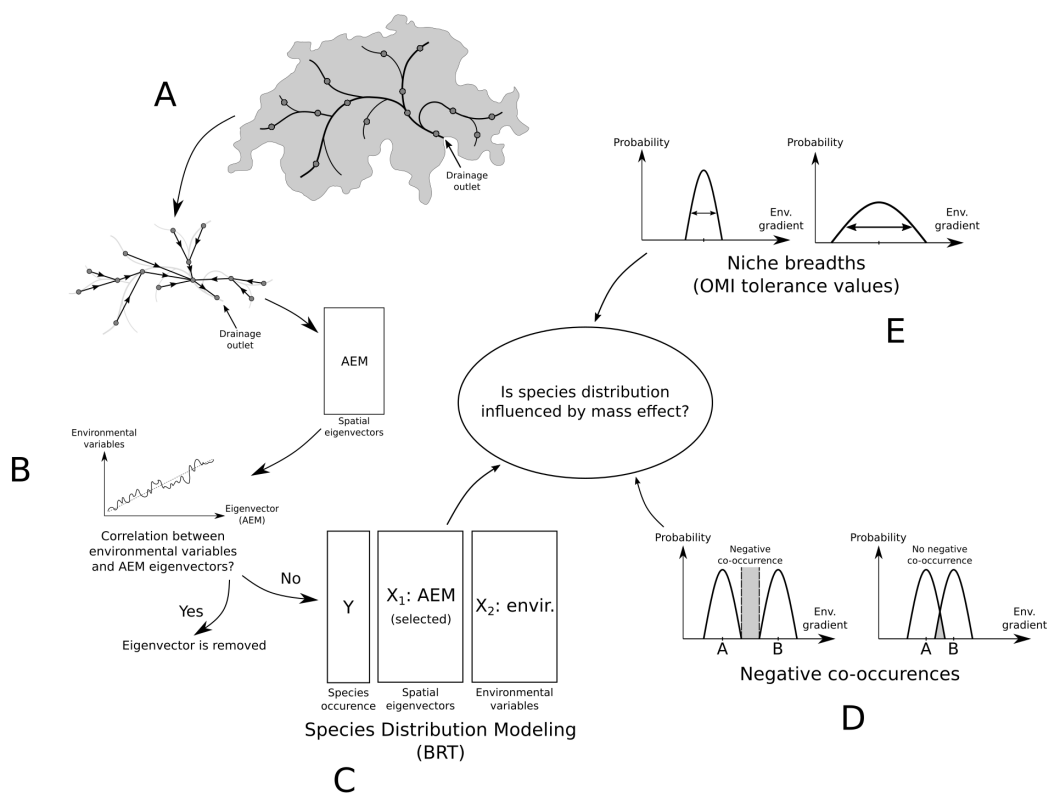


FIGURE 4.2 – Schematic representation of the global procedure for assessing the influence of mass effect on species distribution. We combined three different approaches to detect species influenced by mass effect. (A) We used the AEM eigenvectors not correlated to environmental variables (B), and environmental variables to measure the relative influence of mass effect and environment in structuring species distribution using BRTs (C). (D) Negative co-occurrences provided informations on the non-overlapping niches between species. (E) OMI approach provided tolerance values, which are proxies of niche breadths.

## Materials and methods

### Dataset description

#### *Diatom and environmental condition datasets*

We analyzed datasets from four major French watersheds (Fig. 4.3), the Garonne, the Seine, the Loire and the Rhone, encompassing benthic diatom samples from 223, 73, 404 and 254 spatially distinct localities. Diatoms were sampled following the standardized protocol NFT 90-354 (AFNOR, 2007), whereby stones were scraped during the low flow period, between the 15th of June and the 15th of August, in 2012 to limit successional changes and disturbance effects. Cells were identified at 1000 x magnification by examining 400 cleaned diatom frustules in permanent slides using Krammer and Lange-Bertalot (1986-1991) and Lange-Bertalot (1995-2015, 2000-2013) as identification references. A taxonomic homogenization at the species level was performed with Omnidia 5.3 software (Lecointe et al., 1993). Diatom data also included biovolumes derived from Omnidia 5.3 software and information on ecological guilds (low profile, high profile and motile, Passy (2007)).

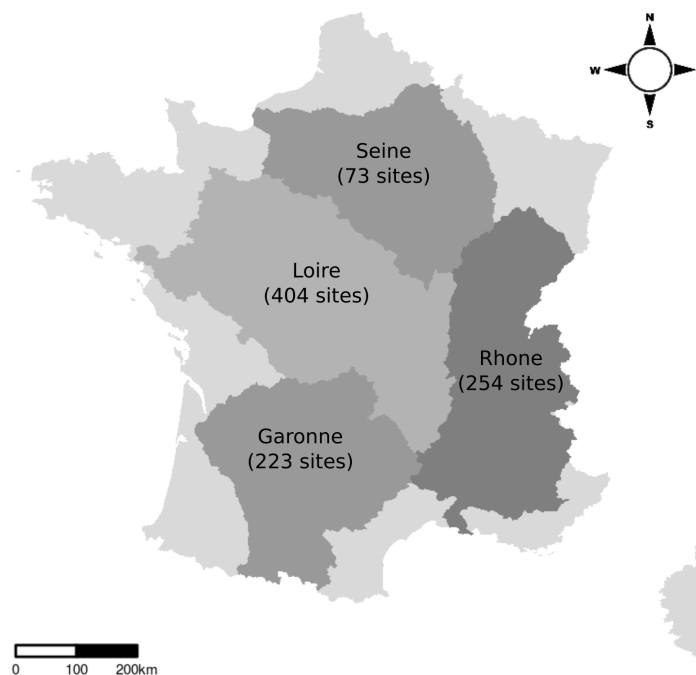


FIGURE 4.3 – Map of the four French watersheds used in this study.

Corresponding physico-chemical data were gathered by the French Water Agencies. Physi-

co-chemical variables included water pH, specific conductance (at 20 °C,  $\text{mS cm}^{-1}$ ), dissolved oxygen saturation (%), biological oxygen demand (mg of  $\text{O}_2$  consumed per liter during 5 days of incubation at 20 °C), concentrations of total phosphorous ( $\text{mg L}^{-1}$  of  $_{15}\text{P}$ ), orthophosphate ( $\text{mg L}^{-1}$  of  $_{15}\text{P}$ ), ammonium ( $\text{mg L}^{-1}$  of  $_{7}\text{N}$ ), nitrate ( $\text{mg L}^{-1}$  of  $_{7}\text{N}$ ) and nitrite ( $\text{mg L}^{-1}$  of  $_{7}\text{N}$ ). We used the median values of all environmental variable measurements obtained during the 30 days before and the 15 days after the diatom sampling date. We also included geographical variables, altitude (m, 1 arc-second) and river slope (%), 1 arc-second), drawn from the Shuttle Radar Topography Mission (Farr et al., 2007) and climatic variables drawn from the WorldClim database (Hijmans et al., 2005). The climatic variables included annual estimates (annual precipitation, annual air temperature), seasonality (temperature annual range and precipitation annual range) and extremes (temperature of the coldest and the warmest month, precipitation of the wet and dry quarters). Climate data covered the years 1950–2000 and average values were used in the analyses.

### *Spatial predictors*

Spatial coordinates provided by the French Water Agencies using Lambert 93 projection (EPSG : 2154) and connectivity derived from Pella et al. (2012) were used to construct AEMs. We tested eigenvectors for spatial autocorrelation using the Moran's I statistic and retained only those with positive spatial autocorrelation. This procedure was conducted separately in each watershed and the generated AEM eigenfunctions were used as spatial explanatory variables.

## **Data analyses**

### *Species distribution modeling*

In order to disentangle the influences of environmental conditions vs. space due to mass effect, we predicted the species occurrences with models integrating all environmental and spatial variables. In all watersheds, we selected species with an occurrence higher than 5 % of all sites, and concerning at least 10 sites, as models do not perform well for rare species (with few occurrences)(Brotons et al., 2004).

Spatial variables correlated to environmental ones (Spearman's correlation,  $\rho > 0.5$ , p-value  $< 0.05$ ) were removed from the analyses to measure 'pure' spatial effect. For each species, a forward selection of environmental (physico-chemical and climatic) and spatial (AEMs) variables was carried out using a cutoff level of  $\alpha = 0.05$  (Blanchet et al., 2008b).

Species distributions were modeled with Boosted Regressions Trees, which is a robust machine learning technique, successfully applied for microorganisms (Pajunen et al., 2016).

BRTs do not require prior data transformation and takes into account the interaction effects between predictors (Elith et al., 2008). BRTs were performed with a Bernoulli distribution using maximum number of 3000 trees, an interaction depth of 3 and a learning rate of 0.01. For each watershed, we randomly drew 80 % of the samples as a training set leaving the remaining 20 % as a test set.

The observed and predicted occurrences of species were compared by calculating the Area Under the Curve (Fielding and Bell, 1997). Area Under the Curve (AUC) evaluate the ability of a fitted model to discriminate presence/absence by plotting the sensitivity against the specificity and then calculates on the graph the area between the plotted values and a null relationship. Values are bounded between 0.5 and 1.0 with 1.0 indicating a perfect discrimination ability. Given that models show intermediate predictive performance at AUC values  $> 0.7$  and excellent performance at AUC values  $> 0.9$  (Heikkinen et al., 2012), we removed models with AUC values  $< 0.7$ . The sum of the relative influences of environmental predictors represented the pure environmental influence, and the sum of the relative influence of spatial predictors represented the 'pure' spatial influence.

#### *Co-occurrences*

We used the recently developed pairwise approaches (Griffith et al., 2016) to measure negative co-occurrences between each species and all other species within each watershed. Results were expressed in number of co-occurrences, weighted by the number of species in each watershed to make values comparable among watersheds. Values are bounded between 0 and 1.

#### *Outlying Mean Index approach*

Outlying Mean Index (OMI) analysis (Dolédec et al., 2000) is a multivariate co-inertia analysis used to describe species niche by its mean position and breadth along environmental gradients. This analysis has the advantage of making no assumption about the shape of the species response to the environment (unimodal or not) and to give equal weights to all sites, regardless of their species richness. It is then possible to estimate the niche breadth of each species along the studied gradient.

In order to reduce niche breadth (hypervolume of  $n$  dimensions) to a two- or three-dimensional volume, we performed principal components analysis (PCA) on centered and standardized environmental variables. We then simplified the analysis by retaining only the components with eigenvalues greater than those randomly generated by a broken-stick model (Jackson, 1993; Legendre and Legendre, 2012). Selected PCA axes were used in the OMI analysis as environmental gradients. We obtained tolerance values (Tol) which are proxies of niche breadth, Tol increasing with niche breadth. This analysis was conducted separately within each watershed.

## *Species classification*

Results from the four watersheds obtained with BRT, negative co-occurrence and OMI analyses were pooled and used to determine if species distributions were influenced by mass effect. We performed a PCA on centered and standardized variables (relative spatial influence, number of spatial eigenvectors used to explain species distribution in BRTs, number of negative co-occurrences and tolerance values). Again, we retained only the components with eigenvalues greater than those randomly generated by a broken-stick model. The k-means clustering algorithm (Hartigan and Wong, 1979) was used to classify species into homogeneous groups, and the optimal number of clusters was fixed by the elbow method (Kodinariya and Makwana, 2013). We then compared the tolerance values (OMI), the number of negative co-occurrences, the relative spatial influence and the biovolumes according to PCA groups with Kruskal-Wallis tests follow-up to Nemenyi post-hoc tests; and the distribution of species among guilds (high profile, low profile and motile) according to PCA groups with exact binomial tests.

## *Software*

All data analyses were performed with R 3.4.4 (R Core Team, 2018), using the 'vegan', 'adespatial' and 'cooccur' packages (Dray et al., 2018; Griffith et al., 2016; Oksanen et al., 2018).

## **Results**

The number of species present in each watershed varied between 363 in the smallest one, the Seine, and 609 in the largest one, the Loire. For each watershed, approximately 25 % of the species were selected and tested with BRTs, and approximately 33 % of these species had AUC higher than 0.7 (Table 4.1). The relative spatial influences ranged between 0 and 100 %, with a mean value of 54.8 %. The number of spatial eigenvectors used to explain species distribution ranged from 0 to 29, with a mean number of 6.8. Negative co-occurrences analysis generated values ranging from 0 to 0.26, and OMI analysis provided Tol values ranging from 0.42 to 39.70 across watersheds (see Appendix S1).

|         | Total number of species | Number of species tested with BRT | Number of species with AUC values >0.7 |
|---------|-------------------------|-----------------------------------|----------------------------------------|
| Garonne | 506                     | 132                               | 44                                     |
| Seine   | 363                     | 80                                | 8                                      |
| Loire   | 609                     | 136                               | 56                                     |
| Rhone   | 462                     | 122                               | 39                                     |

TABLE 4.1 – Species information per watershed.

The relative spatial influence, the number of spatial eigenvectors, the number of negative co-occurrences and the Tol values were centered and standardized to perform a second PCA. With the Elbow method, we classified species into three homogeneous groups (Fig. 4.4).

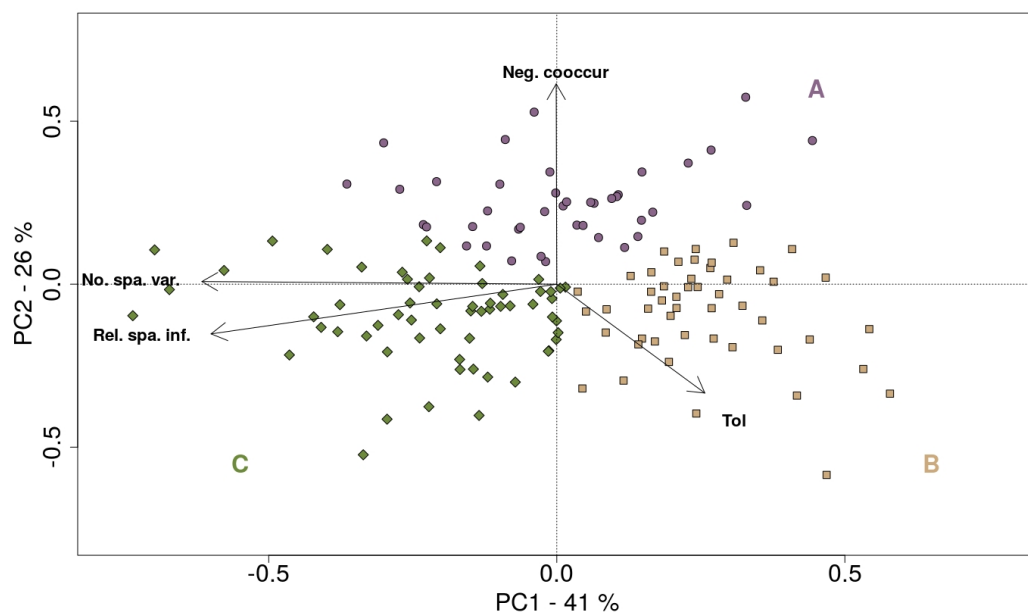


FIGURE 4.4 – Principal component analysis of relative spatial influence (Rel. spa. Inf.), number of spatial eigenvectors used to explain species distribution (No. spa. var.), number of negative co-occurrences and tolerance values of OMI analyses (Tol) for the 147 species observed in the studied watersheds.

Kruskal-Wallis and Nemenyi tests revealed significant differences among the three groups. Group A was characterized by a high number of negative co-occurrences, a medium spatial influence and low tolerance values, group B was characterized by high tolerance values, low number of negative co-occurrences and low spatial influence, and group C was characterized by high spatial influence, low tolerance values and low number of negative co-occurrences (Fig. 4.5). No significant differences between the groups were detected in terms of biovolume or guilds composition (Fig. 4.6).

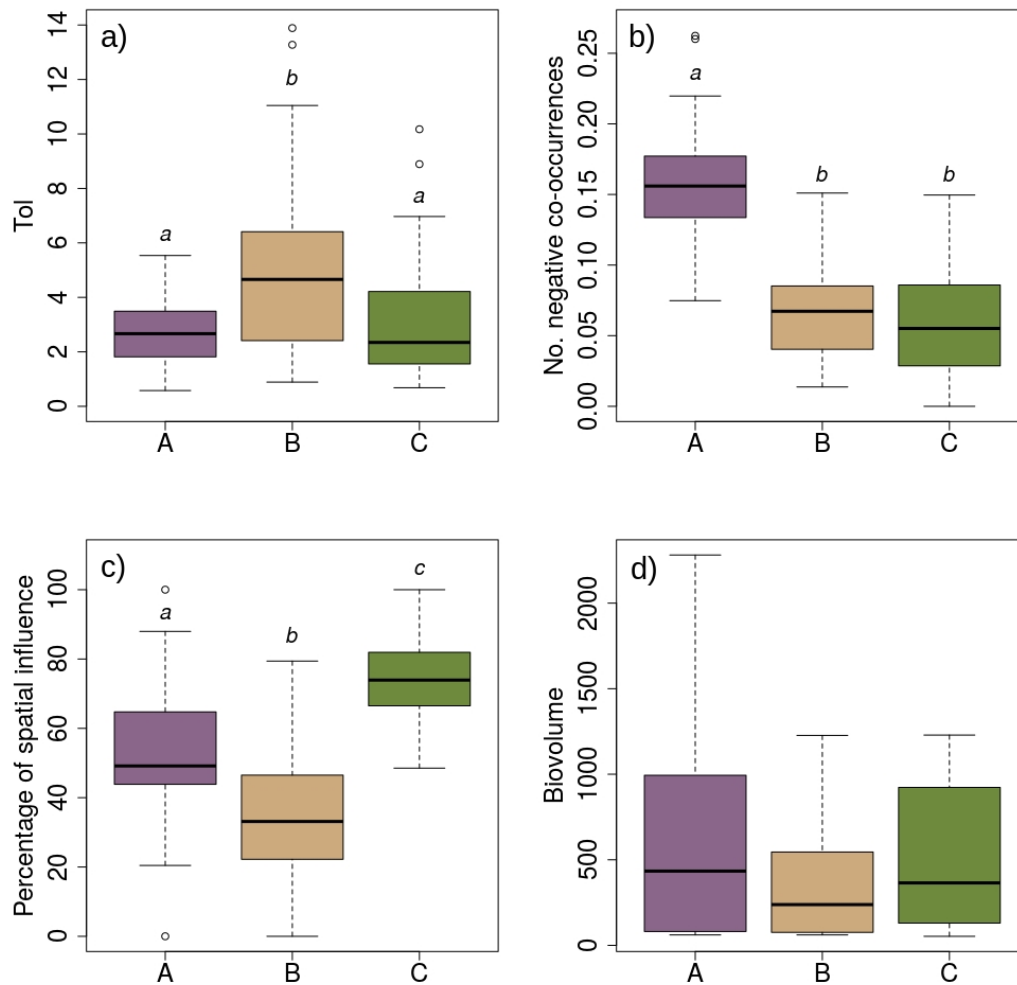


FIGURE 4.5 – Boxplots of (a) the tolerance values (Tol), (b) the number of negative co-occurrences, (c) the relative spatial influence and (d) biovolumes ( $p = 0.39$ ) across the 3 PCA groups. Differences between PCA groups were tested with Kruskal-Wallis tests followed-up by Nemenyi post-hoc tests. Letters represent significant differences between PCA groups, *i.e.*  $p < 0.001$ .

In relation to our conceptual framework (Fig. 4.1), species in PCA group C were suspected to encompass species influenced by mass effect.

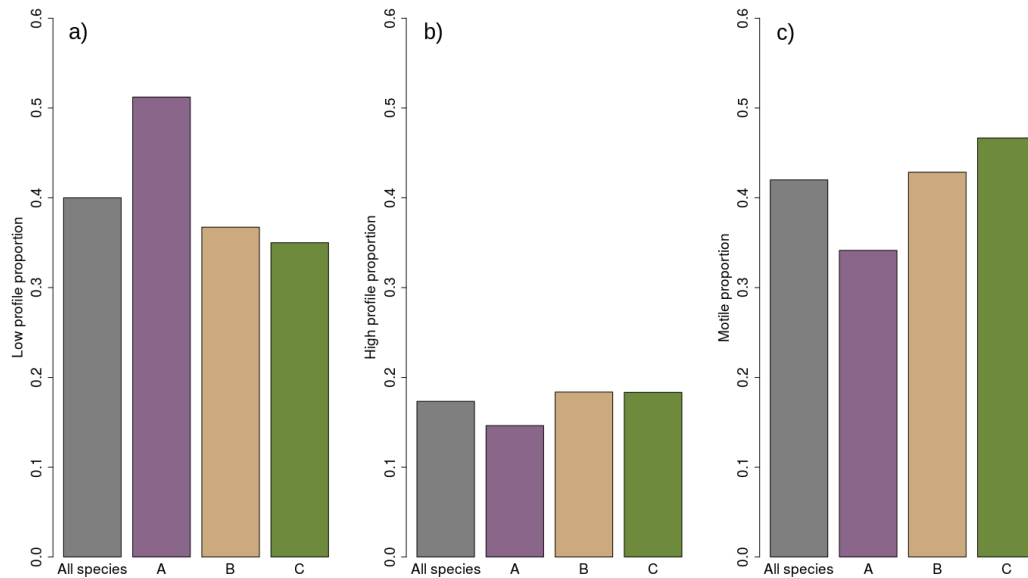


FIGURE 4.6 – Barplots of the proportion of (a) low profile species, (b) high profile species and (c) motile species for all species and the three PCA groups. Exact binomial tests revealed no significant differences between groups.

## Discussion

Here, we developed an innovative and robust method for the detection of species influenced by mass effect. We tested it on stream diatoms and classified a number of species as highly influenced by this process, suggesting that they should be treated with caution when designing river ecological assessment protocols. We discuss the implications of these findings below.

Asymmetric eigenvector maps (AEM) provided an adequate spatial framework taking into account the among-site connectivity and the particularities of river networks. Such considerations, too often ignored in metacommunity models (Brown and Swan, 2010; Cottenie et al., 2003), are essential for a better understanding of the spatial patterns of passive dispersers in river systems (Liu et al., 2013). Further complication in the interpretation of spatial structure arises from the lack of rigorous methodology to distinguish mass effect from limited dispersal. As discussed, currently the two processes are assumed to act at different spatial scales (*i.e.* local scale for mass effect and large scale for dispersal limitation). This, however, may not always be the case, given that physical barriers (*e.g.* moun-

tains or dams) may induce dispersal limitation at small scale (Bottin et al., 2014; Dong et al., 2016). Here, we overcame this problem by developing a framework combining both spatial influence and co-occurrence patterns.

Co-occurrence patterns already provided significant insights into the processes shaping communities. For example, Larsen and Ormerod (2014) reported for stream macroinvertebrates a shift from negative co-occurrence patterns due to biotic interaction (competition, predation) under natural conditions to random species associations due to habitat segregation under anthropogenically disturbed conditions (agriculture). In stream diatom communities, Bottin et al. (2016) found that co-occurrence patterns departed from both random and environmentally predicted structure, and tentatively related these patterns to mass effect. However these authors emphasized the need for further investigations to confirm its importance in structuring co-occurrence patterns. Our approach linked mass effect and co-occurrence patterns and allowed us to identify a group of species with strong correlations with AEM spatial eigenvectors and weak negative co-occurrence patterns (group C), which we consider to be influenced by mass effect. Notably, most species in group C were mesotrophic to eutrophic, typical of not nutrient limited rivers (Van Dam et al., 1994). The capacity to produce high population densities has already been related to nutrient levels as low level of phosphorus (Bothwell, 1989; Chételat et al., 1999) and nitrogen (Stevenson et al., 2006) limiting diatom community growth. In such conditions, the result stating that mesotrophic or eutrophic species are preferably influenced by mass effect seems totally consistent, and in line with Goldenberg Vilar et al. (2014), whose study concludes that in mesotrophic environments, purely spatial processes play a prominent role in structuring diatom communities.

We tested our framework on diatom guilds that are really useful in ecological assessments of human-impacted ecosystems (Passy, 2007), but we did not find any differences between guilds in terms of mass effect. Our results thus disagree with Jamoneau et al. (2018) who suggested that motile species may be affected by dispersal processes. Instead, it became evident that across all guilds there were species driven by either environmental factors or mass effect. We reached a similar conclusion with respect to biovolume, which did not vary across groups. Strong relationships between body size and the strength of the spatial patterns in many metacommunities (12 groups in cited study ranging from bacteria to fish) suggested that dispersal-driven processes were more important for large than small organisms (De Bie et al., 2012). This does not appear to be the case for passive dispersers (Jenkins et al., 2007), including micro-organisms (this study). Indeed, Astorga et al. (2012), working with stream diatoms, macroinvertebrates and bryophytes, underlined that the relative importance of niche vs dispersal-driven processes was not merely related to organism size and other traits (*e.g.* life-history or dispersal-related) can complicated this relationship. We suggest that the sensitivity of species towards mass effect should represent a new ecological trait to be considered for fundamental and applied issues concerning ecology and water quality assessment.

Almost all of the species identified here as influenced by mass effect are contributing to the

calculation of different diatom-based indices such as the Biological Diatom Index (BDI), the standardized method routinely used in France for stream water quality assessment (Coste et al., 2009) or the Specific Pollution-sensitivity Index (SPI), an other method used to detect physicochemical pressures (Coste, 1982). In light of our results, such species then should be treated with caution when assigning ecological status classes to water bodies, and our methodology should also be applied to other taxa-based indices. For example, species influenced by mass effect should have lower weight in index calculation or simply removed. In fact, the presence of species in unsuitable habitats due to mass effect may mask the effect of local environmental conditions (Heino et al., 2015; Leibold et al., 2004) and result in potentially erroneous evaluation of water quality. Thus, we advocate for a better integration of spatial mechanisms in bio-assessment in order to facilitate the development of more effective water quality assessment indices, incorporating both environmental and dispersal processes.

## References

- AFNOR, 2007. Norme NFT 90-354 - Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD).
- Amarasekare, P., Nisbet, R.M., 2001. Spatial Heterogeneity, Source-Sink Dynamics, and the Local Coexistence of Competing Species. *The American Naturalist* 158, 572–584.
- Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R., Muotka, T., 2012. Distance decay of similarity in freshwater communities : do macro- and microorganisms follow the same rules ? *Global Ecology and Biogeography* 21, 365–375.
- Bell, G., Lechowicz, M.J., Waterway, M.J., 2006. The Comparative Evidence Relating to Functional and Neutral Interpretations of Biological Communities. *Ecology* 87, 1378–1386.
- Bennett, J.R., Cumming, B.F., Ginn, B.K., Smol, J.P., 2010. Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography* 19, 724–732.
- Blanchet, G., Legendre, P., Borcard, D., 2008a. Modelling directional spatial processes in ecological data. *Ecological Modelling* 215, 325–336.
- Blanchet, G.F., Legendre, P., Borcard, D., 2008b. Forward Selection of Explanatory Variables. *Ecology* 89, 2623–2632.
- Bothwell, M.L., 1989. Phosphorus-Limited Growth Dynamics of Lotic Periphytic Diatom Communities : Areal Biomass and Cellular Growth Rate Responses. *Canadian Journal of Fisheries and Aquatic Sciences* 46, 1293–1301.
- Bottin, M., Soininen, J., Alard, D., Rosebery, J., 2016. Diatom Cooccurrence Shows Less Segregation than Predicted from Niche Modeling. *PLOS ONE* 11, e0154581.
- Bottin, M., Soininen, J., Ferrol, M., Tison-Rosebery, J., 2014. Do spatial patterns of benthic diatom assemblages vary across regions and years ? *Freshwater Science* 33, 402–416.
- Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H., 2004. Presence-absence versus

- presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27, 437–448.
- Brown, B.L., Swan, C.M., 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology* 79, 571–580.
- Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O., Patrick, C., 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society* 30, 310–327.
- Chase, J.M., 2014. Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science* 25, 319–322.
- Chase, J.M., Amarasekare, P., Cottenie, K., Gonzalez, A., Holt, R.D., Holyoak, M., Hoopes, M.F., Leibold, M.A., Loreau, M., Mouquet, N., Shurin, J.B., Tilman, D., 2005. Competing theories for competitive metacommunities. *Metacommunities : spatial dynamics and ecological communities* , 335–354.
- Chételat, J., Pick, F.R., Morin, A., Hamilton, P.B., 1999. Periphyton biomass and community composition in rivers of different nutrient status. *Canadian Journal of Fisheries and Aquatic Sciences* 56, 560–569.
- Coste, M., 1982. Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse .
- Coste, M., Boutry, S., Tison-Rosebery, J., Delmas, F., 2009. Improvements of the Biological Diatom Index (BDI) : Description and efficiency of the new version (BDI-2006). *Ecological Indicators* 9, 621–650.
- Cottenie, K., Michels, E., Nuytten, N., De Meester, L., 2003. Zooplankton Metacommunity Structure : Regional Vs. Local Processes in Highly Interconnected Ponds. *Ecology* 84, 991–1000.
- De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W., Declerck, S.A.J., 2012. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters* 15, 740–747.
- Dolédec, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche Separation in Community Analysis : A New Method. *Ecology* 81, 2914–2927.
- Dong, X., Li, B., He, F., Gu, Y., Sun, M., Zhang, H., Tan, L., Xiao, W., Liu, S., Cai, Q., 2016.

- Flow directionality, mountain barriers and functional traits determine diatom metacom-  
munity structuring of high mountain streams. *Scientific Reports* 6.
- Dray, S., Blanchet, G., Borcard, D., Clappe, S., Jombart, G.G.T., Larocque, G., Legendre,  
P., Madi, N., Wagner, H.H., 2018. *adespatial* : Moran's eigenvector maps and related  
methods for the spatial multiscale analysis of ecological communities. R package ver-  
sion 0.1-1.
- Ehrlén, J., Eriksson, O., 2000. Dispersal Limitation and Patch Occupancy in Forest Herbs.  
*Ecology* 81, 1667–1674.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees.  
*Journal of Animal Ecology* 77, 802–813.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller,  
M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner,  
M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission.  
*Reviews of Geophysics* 45.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors  
in conservation presence/absence models. *Environmental Conservation* 24, 38–49.
- Finlay, B.J., 2002. Global Dispersal of Free-Living Microbial Eukaryote Species. *Science*  
296, 1061–1063.
- Goldenberg Vilar, A., van Dam, H., van Loon, E.E., Vonk, J.A., van Der Geest, H.G.,  
Admiraal, W., 2014. Eutrophication decreases distance decay of similarity in diatom  
communities. *Freshwater Biology* 59, 1522–1531.
- Grant, E.H.C., Lowe, W.H., Fagan, W.F., 2007. Living in the branches : population dyna-  
mics and ecological processes in dendritic networks. *Ecology Letters* 10, 165–175.
- Griffith, D.M., Veech, J.A., Marsh, C.J., 2016. *Cooccur* : probabilistic species co-  
occurrence analysis in R. *J Stat Softw* 69, 1–17.
- Göthe, E., Angeler, D.G., Gottschalk, S., Löfgren, S., Sandin, L., 2013. The Influence  
of Environmental, Biotic and Spatial Factors on Diatom Metacomunity Structure in  
Swedish Headwater Streams. *PLOS ONE* 8, e72237.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136 : A K-Means Clustering Algorithm.  
*Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100–108.
- Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species  
distribution models come at the expense of transferability ? *Ecography* 35, 276–288.

- 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. *Freshwater Biology* 60, 845–869.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology* 25, 1965–1978.
- Holyoak, M., Leibold, M.A., Holt, R.D., 2005. *Metacommunities : Spatial Dynamics and Ecological Communities*. University of Chicago Press.
- Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis : A Comparison of Heuristical and Statistical Approaches. *Ecology* 74, 2204–2214.
- Jamoneau, A., Passy, S.I., Soininen, J., Leboucher, T., Tison-Rosebery, J., 2018. Beta diversity of diatom species and ecological guilds : Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology* 63, 62–73.
- Jenkins, D.G., Brescacin, C.R., Duxbury, C.V., Elliott, J.A., Evans, J.A., Grablow, K.R., Hillegass, M., Lyon, B.N., Metzger, G.A., Olandese, M.L., Pepe, D., Silvers, G.A., Suresch, H.N., Thompson, T.N., Trexler, C.M., Williams, G.E., Williams, N.C., Williams, S.E., 2007. Does size matter for dispersal distance ? *Global Ecology and Biogeography* 16, 415–425.
- Kodinariya, T.M., Makwana, P.R., 2013. Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies* 1, 90–95.
- Landeiro, V.L., Magnusson, W.E., Melo, A.S., Espírito-Santo, H.M.V., Bini, L.M., 2011. Spatial eigenfunction analyses in stream networks : do watercourse and overland distances produce different results ? *Freshwater Biology* 56, 1184–1192.
- Larsen, S., Ormerod, S.J., 2014. Anthropogenic modification disrupts species co-occurrence in stream invertebrates. *Global Change Biology* 20, 51–60.
- Lecointe, C., Coste, M., Prygiel, J., 1993. Omnidia : software for taxonomy, calculation of diatom indices and inventories management. *Hydrobiologia* 269-270, 509–513.
- Legendre, P., Legendre, L.F.J., 2012. *Numerical Ecology*. Elsevier.
- 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., Gonzalez, A., 2004. The metacommunity concept : a framework for multi-scale community ecology. *Ecology Letters* 7, 601–613.
- Liu, J., Soininen, J., Han, B.P., Declerck, S.A.J., 2013. Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of Biogeography* 40, 2238–2248.
- MacArthur, R.H., Wilson, E.O., 1967. The theory of island biogeography. volume 1. Princeton university press.
- Mouquet, N., Loreau, M., 2003. Community Patterns in Source Sink Metacommunities. *The American Naturalist* 162, 544–557.
- Ng, I.S.Y., Carr, C.M., Cottenie, K., 2009. Hierarchical zooplankton metacommunities : distinguishing between high and limiting dispersal mechanisms. *Hydrobiologia* 619, 133–143.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. *vegan : Community ecology package*. R package version , 117–118.
- Pajunen, V., Luoto, M., Soininen, J., 2016. Climate is an important driver for stream diatom distributions. *Global Ecology and Biogeography* 25, 198–206.
- Palmer, M.A., Allan, J.D., Butman, C.A., 1996. Dispersal as a regional process affecting the local dynamics of marine and stream benthic invertebrates. *Trends in ecology & evolution* 11, 322–326.
- Passy, S.I., 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquatic Botany* 86, 171–178.
- Pella, H., Lejot, J., Lamouroux, N., Snelder, T., 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. *Géomorphologie : relief, processus, environnement* 18, 317–336.
- R Core Team, 2018. *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Shmida, A., Wilson, M.V., 1985. Biological Determinants of Species Diversity. *Journal of Biogeography* 12, 1–20.
- Soininen, J., 2007. Environmental and Spatial Control of Freshwater Diatoms - a Review. *Diatom Research* 22, 473–490.

- Soininen, J., Jamoneau, A., Rosebery, J., Passy, S.I., 2016. Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography* 25, 940–950.
- Soininen, J., Lennon, J.J., Hillebrand, H., 2007. A Multivariate Analysis of Beta Diversity across Organisms and Environments. *Ecology* 88, 2830–2838.
- Stevenson, R.J., Rier, S.T., Riseng, C.M., Schultz, R.E., Wiley, M.J., 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria, in : *Advances in Algal Biology : A Commemoration of the Work of Rex Lowe*. Springer, Dordrecht. *Developments in Hydrobiology*, pp. 149–165.
- Svenning, J.C., Normand, S., Skov, F., 2008. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography* 31, 316–326.
- Tuomisto, H., Ruokolainen, K., Yli-Halla, M., 2003. Dispersal, Environment, and Floristic Variation of Western Amazonian Forests. *Science* 299, 241–244.
- Van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Netherlands Journal of Aquatic Ecology* 28, 117–133.
- Vanormelingen, P., Verleyen, E., Vyverman, W., 2008. The diversity and distribution of diatoms : from cosmopolitanism to narrow endemism. *Biodiversity and Conservation* 17, 393–405.
- Verleyen, E., Vyverman, W., Sterken, M., Hodgson, D.A., De Wever, A., Juggins, S., Van de Vijver, B., Jones, V.J., Vanormelingen, P., Roberts, D., Flower, R., Kilroy, C., Souffreau, C., Sabbe, K., 2009. The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos* 118, 1239–1249.
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Vijver, B.V., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P., Wever, A.D., 2007. Historical Processes Constrain Patterns in Global Diatom Diversity. *Ecology* 88, 1924–1931.

## Acknowledgements

We thank Katrina Pound and Joseph Mruzek for helpful comments during this study. We thank all French Water Agencies for data sharing. We gratefully acknowledge financial

support from AFB (French Agency for Biodiversity), and the French National Research Agency (ANR) in the frame of the Investments for the future Programme, within the Cluster of Excellence COTE (ANR-10-LABX-45).

## Appendix 1 : Species classification

### Garonne – Group A

| Species                                                    | Environmental influence (%) | Number of mental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occur. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|------------------------------------------------------------|-----------------------------|----------------------------------|-----------------------|-----------------------------------|----------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Achnanthyrium rivulare</i><br>Potapova & Ponader        | 47.2                        | 6                                | 52.8                  | 6                                 | 0.16                                   | 2.26            | 73                          | low    | A     |
| <i>Achnanthyrium subudsonis</i><br>(Hustedt) Kobayasi      | 55.1                        | 4                                | 44.9                  | 19                                | 0.18                                   | 0.58            | 66                          | low    | A     |
| <i>Achnanthyrium subatomus</i><br>(Hustedt) Lange-Bertalot | 42.7                        | 1                                | 57.3                  | 2                                 | 0.18                                   | 3.98            | 61                          | low    | A     |
| <i>Adafia lange-bertalotii</i><br>Monnier & Ector          | 55.0                        | 2                                | 45.0                  | 1                                 | 0.13                                   | 1.81            | -                           | motile | A     |
| <i>Amphora pediculus</i> (Kutzing)<br>Grunow               | 50.9                        | 7                                | 49.1                  | 8                                 | 0.19                                   | 2.96            | 72                          | low    | A     |
| <i>Cocconeis euglypta</i><br>Ehrenberg                     | 12.0                        | 2                                | 88.0                  | 7                                 | 0.14                                   | 1.72            | 2533                        | low    | A     |
| <i>Cocconeis placentula</i><br>Ehrenberg                   | 78.4                        | 2                                | 21.6                  | 1                                 | 0.26                                   | 4.04            | 2871                        | low    | A     |
| <i>Cocconeis pseudolineata</i><br>(Geitler) Lange-Bertalot | 29.4                        | 1                                | 70.6                  | 4                                 | 0.17                                   | 3.64            | 1223                        | low    | A     |
| <i>Diploneis oculata</i> (Brebisson)<br>Cleve              | 54.9                        | 2                                | 45.1                  | 3                                 | 0.09                                   | 0.65            | 1022                        | low    | A     |
| <i>Encyonema minutum</i> (Hilse<br>in Rabh.) Mann          | 52.0                        | 2                                | 48.0                  | 4                                 | 0.22                                   | 5.54            | 213                         | high   | A     |
| <i>Fallacia mitis</i> (Hustedt) Mann                       | 60.7                        | 2                                | 39.3                  | 2                                 | 0.10                                   | 1.19            | -                           | motile | A     |
| <i>Fallacia subhamulata</i><br>(Grunow in Heurck) Mann     | 63.8                        | 2                                | 36.2                  | 2                                 | 0.13                                   | 1.58            | 238                         | motile | A     |
| <i>Navicula tripunctata</i> (Muller)<br>Bory               | 72.1                        | 4                                | 27.9                  | 2                                 | 0.21                                   | 3.95            | 966                         | motile | A     |
| <i>Nitzschia inconspicua</i><br>Grunow                     | 23.7                        | 2                                | 76.3                  | 4                                 | 0.15                                   | 3.18            | 89                          | motile | A     |
| <i>Rhoicosphenia abbreviata</i><br>(Agardh) Lange-Bertalot | 14.3                        | 2                                | 85.7                  | 5                                 | 0.15                                   | 3.88            | 568                         | high   | A     |
| <i>Encyonema silesiacum</i><br>(Bleisch in Rabh.) Mann     | 0.0                         | 0                                | 100.0                 | 5                                 | 0.16                                   | 2.80            | 821                         | high   | A     |

## Garonne – Group B

| Species                                                                     | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild   | Group |
|-----------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|---------|-------|
| <i>Amphora copulata</i> (Kützing)                                           | 56.7                        | 2                                       | 43.3                  | 2                                 | 0.07                                 | 1.59            | 20096                       | low     | B     |
| Schoeman & Archibald                                                        |                             |                                         |                       |                                   |                                      |                 |                             |         |       |
| <i>Achnanthyrium minutissimum</i> (Kützing) Czarnecki                       | 52.5                        | 3                                       | 47.5                  | 1                                 | 0.09                                 | 4.05            | 76                          | low     | B     |
| <i>Achnanthyrium saprophilum</i> (Kobayasi & Mayama) Round & Bukhtiyarova   | 65.8                        | 8                                       | 34.2                  | 2                                 | 0.02                                 | 13.89           | 66                          | low     | B     |
| <i>Achnanthyrium subatomoides</i> (Hustedt) Monnier, Lange-Bertalot & Ector | 83.6                        | 5                                       | 16.4                  | 2                                 | 0.10                                 | 5.85            | 118                         | low     | B     |
| <i>Caloneis bacillum</i> (Grunow) Cleve                                     | 50.2                        | 4                                       | 49.8                  | 2                                 | 0.11                                 | 4.66            | 523                         | motile  | B     |
| <i>Diatoma mesodon</i> (Ehrenberg) Kützing                                  | 68.8                        | 2                                       | 31.2                  | 2                                 | 0.07                                 | 4.73            | 1143                        | high    | B     |
| <i>Hannaea arcus</i> (Ehrenberg) Patrick                                    | 57.9                        | 3                                       | 42.1                  | 3                                 | 0.09                                 | 5.69            | -                           | low     | B     |
| <i>Navicula difficillimoides</i> Hustedt                                    | 85.4                        | 3                                       | 14.6                  | 1                                 | 0.04                                 | 1.00            | -                           | motile  | B     |
| <i>Pseudofallacia monoculata</i> (Hustedt) Liu Kociolek & Wang              | 43.1                        | 2                                       | 56.9                  | 2                                 | 0.05                                 | 5.34            | -                           | Unknown | B     |
| <i>Fallacia subulcidula</i> (Hustedt) Mann                                  | 84.6                        | 3                                       | 15.4                  | 1                                 | 0.09                                 | 1.67            | 66                          | motile  | B     |
| <i>Geissleria acceptata</i> (Hustedt) Lange-Bertalot & Meizelstin           | 75.0                        | 7                                       | 25.0                  | 2                                 | 0.08                                 | 5.55            | 112                         | motile  | B     |
| <i>Gomphonema elegans</i> (Reichardt & Lange-Bertalot) Monnier & Ector      | 20.6                        | 1                                       | 79.4                  | 3                                 | 0.09                                 | 10.52           | 360                         | high    | B     |
| <i>Gomphonema exilissimum</i> (Grunow) Lange-Bertalot & Reichardt           | 75.9                        | 4                                       | 24.1                  | 1                                 | 0.02                                 | 8.36            | 431                         | high    | B     |
| <i>Gomphonema olivaceum</i> (Hornemann) Brubisson                           | 70.2                        | 2                                       | 29.8                  | 3                                 | 0.04                                 | 0.92            | 510                         | high    | B     |

| Species                                                | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|--------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Karayevia oblongella</i> (Oestrup)                  | 100.0                       | 9                                       | 0.0                   | 0                                 | 0.06                                 | 7.07            | 220                         | low    | B     |
| <i>Meridion circulare</i> (Greville) Agardh            | 84.3                        | 3                                       | 15.7                  | 1                                 | 0.04                                 | 3.84            | 671                         | high   | B     |
| <i>Tryblionella kuetzingii</i> Alvarez-Blanco & Blanco | 63.3                        | 1                                       | 36.7                  | 3                                 | 0.04                                 | 2.32            | -                           | motile | B     |
| <i>Navicula rhynchocephala</i> Kutzing                 | 58.4                        | 1                                       | 41.6                  | 1                                 | 0.04                                 | 4.19            | 992                         | motile | B     |
| <i>Navicula veneta</i> Kutzing                         | 77.0                        | 4                                       | 23.0                  | 2                                 | 0.06                                 | 7.69            | 279                         | motile | B     |
| <i>Nitzschia angustatula</i> Lange-Bertalot            | 77.8                        | 5                                       | 22.2                  | 2                                 | 0.04                                 | 0.89            | 176                         | motile | B     |

## Garonne – Group C

| Species                                              | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Diatoma vulgare</i> Bory                          | 16.0                        | 2                                       | 84.0                  | 14                                | 0.03                                 | 1.38            | 3033                        | high   | C     |
| <i>Fragilaria vaucheriae</i> (Kützinger) Petersen    | 24.2                        | 2                                       | 75.8                  | 9                                 | 0.12                                 | 5.18            | -                           | high   | C     |
| <i>Fallacia pygmaea</i> (Kützinger) Stickle & Mann   | 30.9                        | 1                                       | 69.1                  | 2                                 | 0.04                                 | 2.18            | 1229                        | motile | C     |
| <i>Gomphonema pumilum</i> Reichardt & Lange-Bertalot | 36.0                        | 4                                       | 64.0                  | 5                                 | 0.10                                 | 4.21            | 270                         | high   | C     |
| <i>Navicula antonii</i> Lange-Bertalot               | 19.5                        | 1                                       | 80.5                  | 9                                 | 0.08                                 | 2.92            | 434                         | motile | C     |
| <i>Nitzschia fonticola</i> Grunow                    | 7.8                         | 1                                       | 92.2                  | 13                                | 0.08                                 | 2.49            | 344                         | motile | C     |
| <i>Nitzschia costei</i> Tudesque, Rimet & Ector      | 6.7                         | 2                                       | 93.3                  | 17                                | 0.06                                 | 10.17           | 170                         | motile | C     |
| <i>Simonsenia delognei</i> Lange-Bertalot            | 21.6                        | 1                                       | 78.4                  | 4                                 | 0.09                                 | 0.68            | 53                          | motile | C     |

## Seine – Group A, B & C

| Species                                                                      | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|------------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Achnanthidium rivulare</i> Potapova & Ponader                             | 100.0                       | 6                                       | 0.0                   | 0                                 | 0.19                                 | 2.90            | 73                          | low    | A     |
| <i>Achnanthidium minutissimum</i> (Kützing) Czarnecki                        | 100.0                       | 3                                       | 0.0                   | 0                                 | 0.01                                 | 6.62            | 76                          | low    | B     |
| <i>Cocconeis placentula</i> Ehrenberg                                        | 59.4                        | 2                                       | 40.6                  | 1                                 | 0.10                                 | 4.93            | 2871                        | low    | B     |
| <i>Sellaphora seminulum</i> (Grunow) Mann                                    | 100.0                       | 3                                       | 0.0                   | 0                                 | 0.01                                 | 8.59            | 69                          | motile | B     |
| <i>Platessa conspicua</i> (Mayer) Lange-Bertalot                             | 11.1                        | 1                                       | 88.9                  | 3                                 | 0.03                                 | 4.75            | 130                         | low    | C     |
| <i>Rhoicosphenia abbreviata</i> (Agardh) Lange-Bertalot                      | 20.2                        | 2                                       | 79.8                  | 4                                 | 0.12                                 | 6.33            | 568                         | high   | C     |
| <i>Amphora pediculus</i> (Kützing) Grunow                                    | 0.0                         | 0                                       | 100.0                 | 3                                 | 0.00                                 | 4.67            | 72                          | low    | C     |
| <i>Pseudostaurosira brevistriata</i> (Grunow in Van Heurck) Williams & Round | 0.0                         | 0                                       | 100.0                 | 7                                 | 0.03                                 | 1.86            | 157                         | high   | C     |

## Loire – Group A

| Species                                                                   | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild            | Group  |
|---------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|------------------|--------|
| <i>Amphora pediculus</i> (Kützing)<br>Grunow                              | 52.7                        | 5                                       | 47.3                  | 16                                | 0.17                                 | 2.07            | 72                          | low              | A      |
| <i>Cocconeis euglypta</i><br>Ehrenberg                                    | 35.3                        | 2                                       | 64.7                  | 6                                 | 0.26                                 | 3.49            | 2533                        | low              | A      |
| <i>Cocconeis placentula</i><br>Ehrenberg                                  | 79.6                        | 4                                       | 20.4                  | 1                                 | 0.15                                 | 3.83            | 2871                        | low              | A      |
| <i>Cocconeis pediculus</i><br>Ehrenberg                                   | 35.2                        | 1                                       | 64.8                  | 3                                 | 0.16                                 | 2.67            | 2281                        | low              | A      |
| <i>Gyrosigma sciottense</i><br>(Sullivan & Wormley) Cleve                 | 56.2                        | 3                                       | 43.8                  | 2                                 | 0.09                                 | 1.28            | 2923                        | motile           | A      |
| <i>Mayamaea perinitis</i> (Hustedt)<br>Monnier & Ector                    | 32.0                        | 4                                       | 68.0                  | 10                                | 0.15                                 | 3.64            | 66                          | motile           | A      |
| <i>Navicula cryptotenella</i><br>Lange-Bertalot                           | 56.1                        | 5                                       | 43.9                  | 8                                 | 0.16                                 | 2.95            | 386                         | motile           | A      |
| <i>Navicula tripunctata</i> (Muller)<br>Bory                              | 74.0                        | 4                                       | 26.0                  | 2                                 | 0.17                                 | 1.81            | 966                         | motile           | A      |
| <i>Navicula gregaria</i> Donkin<br><i>Nitzschia inconspicua</i><br>Grunow | 51.2<br>43.0                | 8<br>8                                  | 48.8<br>57.0          | 10<br>21                          | 0.09<br>0.18                         | 2.01<br>2.67    | 485<br>89                   | motile<br>motile | A<br>A |
| <i>Rhoicosphenia abbreviata</i><br>(Agardh) Lange-Bertalot                | 44.7                        | 2                                       | 55.3                  | 4                                 | 0.13                                 | 2.42            | 568                         | high             | A      |
| <i>Reimeria sinuata</i> (Gregory)<br>Kociolek & Stoermer                  | 36.4                        | 2                                       | 63.6                  | 8                                 | 0.15                                 | 4.02            | 406                         | low              | A      |

## Loire – Group B

| Species                                                                      | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occur. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|------------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|----------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Achnanthyidium bioretii</i> (Germain) Edlund                              | 41.9                        | 1                                       | 58.1                  | 2                                 | 0.07                                   | 5.89            | 517                         | low    | B     |
| <i>Achnanthyidium minutissimum</i> (Kützinger) Czarnecki                     | 81.6                        | 5                                       | 18.4                  | 1                                 | 0.07                                   | 7.97            | 76                          | low    | B     |
| <i>Achnanthyidium rivulare</i> Potapova & Ponader                            | 64.4                        | 6                                       | 35.6                  | 4                                 | 0.07                                   | 4.09            | 73                          | low    | B     |
| <i>Achnanthyidium subatomoides</i> (Hustedt) Monnier, Lange-Bertalot & Ector | 58.8                        | 2                                       | 41.2                  | 3                                 | 0.13                                   | 4.56            | 118                         | low    | B     |
| <i>Achnanthyidium subatomus</i> (Hustedt) Lange-Bertalot                     | 53.5                        | 6                                       | 46.5                  | 9                                 | 0.10                                   | 5.36            | 61                          | low    | B     |
| <i>Encyonema minutum</i> (Hilse in Rabh.) Mann                               | 52.4                        | 3                                       | 47.6                  | 3                                 | 0.15                                   | 7.11            | 213                         | high   | B     |
| <i>Fallacia subhamulata</i> (Grunow in Van Heurck) Mann                      | 62.2                        | 4                                       | 37.8                  | 2                                 | 0.07                                   | 3.68            | 238                         | motile | B     |
| <i>Gomphonella calcituga</i> (Lange-Bertalot & Reichardt) Tuji               | 58.0                        | 1                                       | 42.0                  | 1                                 | 0.06                                   | 6.41            | 348                         | high   | B     |
| <i>Meridion circulare</i> (Greville) Agardh                                  | 41.2                        | 2                                       | 58.8                  | 3                                 | 0.06                                   | 11.05           | 671                         | high   | B     |
| <i>Mayamaea ingenua</i> (Hustedt) Lange-Bertalot & Hofmann                   | 76.3                        | 4                                       | 23.7                  | 2                                 | 0.02                                   | 4.51            | -                           | motile | B     |
| <i>Navicula cryptocephala</i> Kützinger                                      | 70.5                        | 3                                       | 29.5                  | 1                                 | 0.08                                   | 2.09            | 431                         | motile | B     |
| <i>Sellaphora seminulum</i> (Grunow) Mann                                    | 100.0                       | 4                                       | 0.0                   | 0                                 | 0.07                                   | 3.86            | 69                          | motile | B     |
| <i>Navicula lanceolata</i> (Agardh) Ehrenberg                                | 73.7                        | 6                                       | 26.3                  | 2                                 | 0.09                                   | 2.41            | 1227                        | motile | B     |
| <i>Navicula rhynchocephala</i> Kützinger                                     | 45.8                        | 2                                       | 54.2                  | 4                                 | 0.07                                   | 4.35            | 992                         | motile | B     |
| <i>Platessa conspicua</i> (Mayer) Lange-Bertalot                             | 48.9                        | 4                                       | 51.1                  | 6                                 | 0.07                                   | 4.09            | 130                         | low    | B     |

# Loire – Group C

| Species                                                                      | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occur. (weigthed) | Tolerance (OMI) | BioVol. (µm <sup>3</sup> ) | Guild  | Group |
|------------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|-----------------------------------------|-----------------|----------------------------|--------|-------|
| <i>Amphora copulata</i> (Kützing) Schoeman & Archibald                       | 23.1                        | 2                                       | 76.9                  | 10                                | 0.08                                    | 1.77            | 20096                      | low    | C     |
| <i>Achnanthyrium eutrophilum</i> (Lange-Bertalot) Lange-Bertalot             | 12.3                        | 2                                       | 87.7                  | 29                                | 0.05                                    | 1.97            | 72                         | low    | C     |
| <i>Achnanthyrium lauenburgianum</i> (Hustedt) Monnier Lange-Bertalot & Ector | 25.7                        | 1                                       | 74.3                  | 6                                 | 0.05                                    | 1.79            | 99                         | low    | C     |
| <i>Achnanthyrium subhudsonis</i> (Hustedt) Kobayasi                          | 34.1                        | 1                                       | 65.9                  | 6                                 | 0.05                                    | 1.95            | 66                         | low    | C     |
| <i>Planorhynchium granum</i> (Hohn & Hellerman) Lange-Bertalot               | 36.9                        | 3                                       | 63.1                  | 9                                 | 0.01                                    | 1.26            | 56                         | low    | C     |
| <i>Cocconeis neothumensis</i> Krammer                                        | 29.7                        | 2                                       | 70.3                  | 10                                | 0.02                                    | 4.21            | 94                         | low    | C     |
| <i>Cocconeis placentula</i> Ehrenberg                                        | 39.8                        | 3                                       | 60.2                  | 5                                 | 0.09                                    | 3.73            | 2963                       | low    | C     |
| <i>Diploneis oculata</i> (Brebisson) Cleve                                   | 32.5                        | 1                                       | 67.5                  | 10                                | 0.02                                    | 1.12            | 1022                       | low    | C     |
| <i>Eolimna minima</i> (Grunow) Lange-Bertalot                                | 28.1                        | 3                                       | 71.9                  | 10                                | 0.02                                    | 3.57            | 88                         | motile | C     |
| <i>Craticula subminuscula</i> (Manguin) Wetzel & Ector                       | 33.7                        | 5                                       | 66.3                  | 20                                | 0.11                                    | 2.41            | -                          | motile | C     |
| <i>Fragilaria vaucheriae</i> (Kützing) Petersen                              | 29.7                        | 2                                       | 70.3                  | 8                                 | 0.07                                    | 3.45            | -                          | high   | C     |
| <i>Fallacia militis</i> (Hustedt) Mann                                       | 18.2                        | 2                                       | 81.8                  | 9                                 | 0.04                                    | 1.37            | -                          | motile | C     |
| <i>Fallacia subcladula</i> (Hustedt) Mann                                    | 48.3                        | 3                                       | 51.7                  | 8                                 | 0.05                                    | 2.28            | 66                         | motile | C     |
| <i>Gomphonema minutum</i> (Agardh)                                           | 8.2                         | 2                                       | 91.8                  | 25                                | 0.08                                    | 1.68            | 318                        | high   | C     |
| <i>Gomphonema pumilum</i> Reichardt & Lange-Bertalot                         | 22.8                        | 3                                       | 77.2                  | 11                                | 0.09                                    | 1.98            | 270                        | high   | C     |
| <i>Nitzschia amphibia</i> Grunow                                             | 25.7                        | 5                                       | 74.3                  | 18                                | 0.13                                    | 4.85            | 334                        | motile | C     |
| <i>Navicula antonii</i> Lange-Bertalot                                       | 31.4                        | 5                                       | 68.6                  | 23                                | 0.12                                    | 1.87            | 434                        | motile | C     |

| Species                                                | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|--------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Nitzschia capitellata</i> Hustedt in Schmidt        | 40.2                        | 3                                       | 59.8                  | 6                                 | 0.02                                 | 3.38            | 549                         | motile | C     |
| <i>Navicula capitatoradiata</i> Germain                | 3.5                         | 1                                       | 96.5                  | 24                                | 0.11                                 | 0.70            | 705                         | motile | C     |
| <i>Navicula germainii</i> Wallace                      | 31.9                        | 3                                       | 68.1                  | 19                                | 0.03                                 | 0.83            | 539                         | motile | C     |
| <i>Nitzschia recta</i> Hantzsch in Rabenhorst          | 34.3                        | 4                                       | 65.7                  | 8                                 | 0.04                                 | 0.97            | 871                         | motile | C     |
| <i>Nitzschia sociabilis</i> Hustedt                    | 39.0                        | 4                                       | 61.0                  | 4                                 | 0.06                                 | 1.82            | 391                         | motile | C     |
| <i>Nitzschia subacicularis</i> Hustedt in Schmidt      | 25.7                        | 4                                       | 74.3                  | 14                                | 0.02                                 | 1.22            | 82                          | motile | C     |
| <i>Nitzschia tubicola</i> Grunow                       | 31.2                        | 1                                       | 68.8                  | 3                                 | 0.04                                 | 3.06            | 336                         | motile | C     |
| <i>Nitzschia costei</i> Tudesque, Rimet & Ector        | 18                          | 3                                       | 82                    | 25                                | 0.11                                 | 3.42            | 170                         | motile | C     |
| <i>Planothidium rostratum</i> (Oestrup) Lange-Bertalot | 17.2                        | 2                                       | 82.8                  | 6                                 | 0.02                                 | 4.24            | 267                         | low    | C     |
| <i>Reimeria uniseriata</i> Sala Guerrero & Ferrario    | 33.3                        | 2                                       | 66.7                  | 13                                | 0.08                                 | 1.11            | 456                         | low    | C     |
| <i>Encyonema lange-bertalotii</i> Krammer              | 0.0                         | 0                                       | 100.0                 | 14                                | 0.02                                 | 1.56            | 1094                        | high   | C     |

## Rhone – Group A

| Species                                                                | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Achnanthyrium pyrenaicum</i> (Hustedt) Kobayasi                     | 34.2                        | 3                                       | 65.8                  | 17                                | 0.20                                 | 4.69            | 106                         | low    | A     |
| <i>Achnanthyrium rivulare</i> Potapova & Ponader                       | 50.2                        | 1                                       | 49.8                  | 6                                 | 0.07                                 | 0.8             | 73                          | low    | A     |
| <i>Achnanthyrium subhudsonis</i> (Hustedt) Kobayasi                    | 75.1                        | 3                                       | 24.9                  | 9                                 | 0.09                                 | 0.7             | 66                          | low    | A     |
| <i>Amphora pediculus</i> (Kützing) Grunow                              | 71.6                        | 8                                       | 28.4                  | 8                                 | 0.15                                 | 3.04            | 72                          | low    | A     |
| <i>Cymbella excisa</i> Kützing                                         | 46.3                        | 6                                       | 53.7                  | 12                                | 0.16                                 | 3.39            | 520                         | low    | A     |
| <i>Cocconeis euglypta</i> Ehrenberg                                    | 33.9                        | 2                                       | 66.1                  | 7                                 | 0.18                                 | 2.24            | 2533                        | low    | A     |
| <i>Cocconeis placentula</i> Ehrenberg                                  | 54.3                        | 5                                       | 45.7                  | 5                                 | 0.15                                 | 2.43            | 2871                        | low    | A     |
| <i>Encyonema minutum</i> (Hilse in Rabh.) Mann                         | 40.1                        | 4                                       | 59.9                  | 8                                 | 0.22                                 | 2.18            | 213                         | high   | A     |
| <i>Gomphonema elegans</i> (Reichardt & Lange-Bertalot) Monnier & Ector | 56.9                        | 4                                       | 43.1                  | 5                                 | 0.16                                 | 3.13            | 360                         | high   | A     |
| <i>Navicula antonii</i> Lange-Bertalot                                 | 42.4                        | 3                                       | 57.6                  | 6                                 | 0.10                                 | 2.29            | 434                         | motile | A     |
| <i>Navicula tripunctata</i> (Muller) Bory                              | 52.6                        | 2                                       | 47.4                  | 5                                 | 0.15                                 | 2.6             | 966                         | motile | A     |
| <i>Navicula gregaria</i> Donkin                                        | 21.2                        | 2                                       | 78.8                  | 6                                 | 0.12                                 | 1.77            | 485                         | motile | A     |

## Rhone – Group B

| Species                                                        | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occ. (weighed) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|----------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|--------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Achnanthydium subatomus</i> (Hustedt) Lange-Bertalot        | 91.4                        | 6                                       | 8.6                   | 2                                 | 0.07                                 | 3.02            | 61                          | low    | B     |
| <i>Caloneis bacillum</i> (Grunow) Cleve                        | 100.0                       | 4                                       | 0.0                   | 0                                 | 0.07                                 | 1.38            | 523                         | motile | B     |
| <i>Caloneis lanceolata</i> (Schulz) Lange-Bertalot & Witkowski | 68.5                        | 4                                       | 31.5                  | 2                                 | 0.02                                 | 1.76            | 626                         | motile | B     |
| <i>Cocconeis placentula</i> Ehrenberg                          | 87                          | 8                                       | 13                    | 5                                 | 0.05                                 | 1.92            | 2533                        | low    | B     |
| <i>Fallacia subhamulata</i> (Grunow in Van Heurck) Mann        | 77.3                        | 4                                       | 22.7                  | 1                                 | 0.07                                 | 1.76            | 238                         | motile | B     |
| <i>Mayamaea perinitis</i> (Hustedt) Monnier & Ector            | 49.4                        | 5                                       | 50.6                  | 5                                 | 0.08                                 | 9.1             | 66                          | motile | B     |
| <i>Tryblionella kuetzingii</i> Alvarez-Blanco & Blanco         | 47.8                        | 1                                       | 52.2                  | 2                                 | 0.02                                 | 1.83            | -                           | motile | B     |
| <i>Navicula veneta</i> Kützing                                 | 54.8                        | 3                                       | 45.2                  | 3                                 | 0.04                                 | 4.8             | 279                         | motile | B     |
| <i>Planorhynchium frequentissimum</i> (Lange-Bertalot)         | 29.6                        | 2                                       | 70.4                  | 11                                | 0.11                                 | 13.27           | 219                         | low    | B     |
| <i>Rhoicosphenia abbreviata</i> (Agardh) Lange-Bertalot        | 66.9                        | 4                                       | 33.1                  | 2                                 | 0.09                                 | 5.74            | 568                         | high   | B     |

## Rhone – Group C

| Species                                                  | Environmental influence (%) | Number of environmental variables (BRT) | Spatial influence (%) | Number of spatial variables (BRT) | Number of negative co-occur. (weighted) | Tolerance (OMI) | BioVol. ( $\mu\text{m}^3$ ) | Guild  | Group |
|----------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------|-----------------------------------|-----------------------------------------|-----------------|-----------------------------|--------|-------|
| <i>Amphora copulata</i> (Kützing)                        | 20.4                        | 1                                       | 79.6                  | 5                                 | 0.06                                    | 1.41            | 20096                       | low    | C     |
| Schoeman & Archibald                                     |                             |                                         |                       |                                   |                                         |                 |                             |        |       |
| <i>Amphora ovalis</i> (Kützing)                          | 23.7                        | 2                                       | 76.3                  | 13                                | 0.04                                    | 1.55            | 52858                       | low    | C     |
| Kützing                                                  |                             |                                         |                       |                                   |                                         |                 |                             |        |       |
| <i>Karayevia ploenensis</i> (Hustedt) Bukhtivarova       | 19.4                        | 1                                       | 80.6                  | 9                                 | 0.03                                    | 1.68            | 331                         | low    | C     |
| <i>Cocconeis pseudolineata</i> (Geitler) Lange-Bertalot  | 49.5                        | 4                                       | 50.5                  | 7                                 | 0.06                                    | 1.56            | 1223                        | low    | C     |
| <i>Cocconeis pediculus</i> Ehrenberg                     | 27.4                        | 1                                       | 72.6                  | 4                                 | 0.04                                    | 4.34            | 2281                        | low    | C     |
| <i>Diatoma moniliformis</i> Kützing                      | 31.4                        | 2                                       | 68.6                  | 4                                 | 0.07                                    | 3.71            | 303                         | high   | C     |
| <i>Encyonopsis minuta</i> Krammer & Reichardt            | 39.4                        | 2                                       | 60.6                  | 13                                | 0.11                                    | 2.49            | 60                          | low    | C     |
| <i>Eolimna minima</i> (Grunow) Lange-Bertalot            | 21.1                        | 3                                       | 78.9                  | 14                                | 0.15                                    | 6.97            | 88                          | motile | C     |
| <i>Craticula subminuscula</i> (Manguin) Wetzel & Ector   | 31.5                        | 3                                       | 68.5                  | 10                                | 0.10                                    | 5.82            | -                           | motile | C     |
| <i>Gomphonema olivaceum</i> (Hornemann) Brubisson        | 27.4                        | 4                                       | 72.6                  | 8                                 | 0.11                                    | 4.98            | 510                         | high   | C     |
| <i>Gyrosigma attenuatum</i> (Kützing) Rabenhorst         | 39.9                        | 1                                       | 60.1                  | 3                                 | 0.02                                    | 1.16            | 11622                       | motile | C     |
| <i>Luticola goeppertiana</i> (Bleich in Rabenhorst) Mann | 13.0                        | 4                                       | 87.0                  | 19                                | 0.00                                    | 8.89            | 925                         | motile | C     |
| <i>Navicula capitatoradiata</i> Germain                  | 26.4                        | 3                                       | 73.6                  | 13                                | 0.06                                    | 2.43            | 705                         | motile | C     |
| <i>Navicula cryptotenella</i> Lange-Bertalot             | 51.5                        | 3                                       | 48.5                  | 9                                 | 0.08                                    | 2.96            | 386                         | motile | C     |
| <i>Nitzschia sociabilis</i> Hustedt                      | 9.2                         | 1                                       | 90.8                  | 11                                | 0.03                                    | 6.72            | 391                         | motile | C     |
| <i>Nitzschia costei</i> Tudesque, Rimet & Ector          | 37.7                        | 1                                       | 62.3                  | 12                                | 0.10                                    | 0.89            | 170                         | motile | C     |
| <i>Sellaphora minima</i> (Grunow) Mann                   | 44.7                        | 7                                       | 55.3                  | 12                                | 0.07                                    | 3.94            | 88                          | motile | C     |



## 5 Influence de l'effet de masse en bio-indication : exemple de l'Indice Biologique Diatomée

Ce chapitre présente un article en cours de relecture auprès des co-auteurs qui sera soumis à la revue *Ecological Indicators*.

### 5.1 Contexte et objectifs de l'étude

Les résultats du chapitre précédent ont permis de mettre en évidence l'impact de l'effet de masse sur la structuration d'une cinquantaine d'espèces, dont 40 rentrent actuellement dans le calcul de l'Indice Biologique Diatomée. Or, ce phénomène n'est actuellement pas du tout considéré dans le calcul de l'indice, et la bio-indication en général est encore largement appréhendée uniquement sous le prisme de la sensibilité ou de la tolérance des espèces à leur environnement local et à sa dégradation par des perturbations anthropiques.

Toutefois, la littérature récente commence à mieux intégrer la dispersion en bio-indication, surtout en rivières où leurs structures dendritiques génèrent une forte dynamique spatiale (Tonkin et al., 2018). Il est notamment largement accepté aujourd'hui que la bio-indication doit être abordée à de petites échelles afin de réduire l'effet du *pool* d'espèces entre différentes zones biogéographiques (Pont et al., 2006; Smucker and Vis, 2011a,b). En revanche, la bio-indication n'intègre actuellement pas l'effet de masse qui permet pourtant à des espèces de survivre en dehors de leur optimum écologique dans des sites qui lui sont *a priori* peu favorables (Mouquet and Loreau, 2003; Shmida and Wilson, 1985), induisant un potentiel biais sur l'évaluation de l'état écologique des sites concernés.

Nous proposons donc ici de décrire concrètement l'impact de l'effet de masse sur la bio-indication, en prenant comme exemple l'IBD (Coste et al., 2009; Lenoir and Coste, 1996) et en comparant deux versions de cet indice : la version originale actuellement appliquée

et une version modifiée où les 40 espèces influencées par l'effet de masse ont été retirées de la liste des espèces indicatrices considérées dans l'indice.

## **5.2 Principaux résultats**

Nous nous sommes basés sur l'ensemble de l'historique disponible dans la base Pandore depuis la mise à jour de l'IBD en 2007 pour réaliser cette étude. Celui-ci regroupe 9487 échantillonnages diatomiques collectés entre 2007 et 2013, répartis sur 3913 sites distincts couvrant l'ensemble du territoire français métropolitain.

La prise en compte de l'effet de masse dans le calcul de l'indice révèle une réponse des notes de l'IBD modifié mieux corrélée aux conditions environnementales locales. La différence entre les deux versions de l'IBD repose essentiellement sur une meilleure discrimination des notes dans les conditions intermédiaires (bon état et état moyen).

Ce travail a permis de démontrer, pour la première fois à notre connaissance, que l'effet de masse avait un fort impact en bio-indication. Sa prise en compte permet en effet d'obtenir une réponse plus fine de l'indice à l'environnement local, conclusion majeure à intégrer lors de l'élaboration de nouvelles méthodes de bio-indication.

# **Taking into account mass effect processes in bioindication allows a more accurate bioassessment of water quality**

**Thibault Leboucher<sup>1\*</sup>, Lucas Mignien<sup>1</sup>, Marie Wach<sup>1</sup>, Sébastien Boutry<sup>1</sup>, Aurélien Jamoneau<sup>1</sup>, Sophia I. Passy<sup>2</sup> and Juliette Tison-Rosebery<sup>1</sup>**

<sup>1</sup> Aquatic Ecosystems and Global Changes Research Unit, IRSTEA, 50 avenue de Verdun, 33612 Cestas, France

<sup>2</sup> Department of Biology, University of Texas at Arlington, Box 19498, Arlington, TX 76019, USA

\* Corresponding author : [thibault.leboucher@irstea.fr](mailto:thibault.leboucher@irstea.fr)

## **Abstract**

Environmental assessment based on biotic indices has long been considered uniquely through the prism of species tolerance and sensitivity towards disturbance, community composition and modifications under anthropogenic threats. However bioassessment would benefit from being considered from a metacommunity perspective integrating dispersal processes, and in particular mass effect. This process is known to be challenging to describe but Leboucher et al. (*in prep.*) partially ruled out this problem for benthic diatoms thanks to the development of an innovative method, leading to a list of 40 diatom species able to produce mass effect patterns. Such species are to be carefully considered when assigning ecological status to water bodies.

Based on these new insights, we propose here to concretely integrate the influence of mass effect in diatom-based assessment methods. More specifically, we aim to generate an alternative version of the Biological Diatom Index (the French diatom-based index used in routine) where species influenced by mass effect were removed from the list of indicator taxa. We analyzed a benthic diatom dataset from France collected between 2007 and 2013, including 9487 samples from 3913 spatially distinct localities and compared the distributions of scores and ecological status assessments obtained with the new and the original BDI versions, and their responses towards local environmental conditions.

Our results indicate a better relationship between the modified BDI version scores and the environmental conditions, comparing to the original BDI scores. These significant differences led to a better discrimination of river ecological status in conditions of moderate anthropogenic pressure, in better line with the nutrient enrichment gradient.

Our results largely support the idea that mass effect induce a potential biased evaluation of water quality. The comparison of the two BDI versions shown in every respect that taking into account this process leads to a more acute assessment of river ecological status. This advocates for an application of our method to any diatom-based indices, and by extension to any biotic index based on active or passive dispersers.

**Keywords :** bio-assessment, community assembly, Biological Diatom Index, diatoms, dispersal, mass effect

## Introduction

Metacommunity ecology assumes that the structure and dynamics of local communities are connected by species dispersal (Leibold et al., 2004). This new framework changed the initial perception of local communities consisting of isolated patches patterned by biotic and abiotic processes to a more global approach of a set of communities connected within a region, and belonging to a common regional species pool. Different conceptual perspectives differ according to the respective influence of species responses towards environmental conditions, and species dispersal abilities. If unlimited dispersal allows species to track favorable biotic and abiotic conditions through species sorting, significant pure spatial structuring of community (*e.g.* mass effect or dispersal limitation) is involved in patch dynamic, mass effect and neutral concepts. If these four paradigms can be theoretically described independently, many studies based on field data in freshwater ecology suggest that their influence seems rather combined, with a consensus on the prevalence of species sorting (Cottenie, 2005). This interplay between processes varies according to the organism groups and to the geographic scale of observation (Johnson et al., 2007), making somewhat difficult to predict how perturbations will alter ecosystems.

Environmental assessment has long been considered uniquely through the prism of species tolerance and sensitivity towards disturbance, community composition and modifications under anthropogenic threats being the foundation of any biotic index. Bioassessment however recently gained to be considered in a metacommunity perspective (Heino et al., 2013) integrating dispersal processes, with concrete implications in theoretical but also applied fields of ecology (Brown et al., 2011). On the first hand, dispersal limitation, which prevents species from reaching suitable habitats (Heino et al., 2015), is now considered in new environmental assessment methods, for freshwater ecosystems in particular, given that their dendritic structure is considered to promote strong spatial dynamics (Tonkin et al., 2018). Approaches are consequently more and more designed for small regional extents, to reduce species pool effects between biogeographical areas (Pont et al., 2006; Smucker and Vis, 2011a,b), and to identify relevant regional scales to design river management programs (Smucker and Vis, 2011b). On the other hand, environmental assessment, to our knowledge, still does not integrate the potential influence of mass effect, which involves the dispersal of species from source locations where local environmental conditions enable high population densities to sinks. Such process allows the persistence of species in unfavorable habitats (Mouquet and Loreau, 2003; Shmida and Wilson, 1985), due to a high influx of colonists may prevent pure species sorting to take place, and may in the same time allow species to locally persist out of their niche (Mouquet and Loreau, 2003). In a bioassessment context, there is thus a need for considering only species not sensitive to mass effect which may then provide more reliable information on ecosystem ecological status (Leboucher et al., *in prep.*).

Diatoms provide a good illustration of this scientific gap. Considered as good indicators

of river water quality (Prygiel et al., 1999), research as long focused exclusively on the relationships between diatom communities and local environmental factors (Potapova and Charles, 2002; Tison et al., 2005). Today a large consensus considers that diatoms exhibit biogeographical patterns (Passy, 2009; Vanormelingen et al., 2008; Vyverman et al., 2007), with a set of evidences towards dispersal limitation signals at regional (Dong et al., 2016) and global scales (Soininen et al., 2016). At the watercourse scale, recent studies tend to advocate for an environmental control of community structure in headwaters, interacting with dispersal driven processes mid- and down-stream (Jamoneau et al., 2018). Although some studies evoked mass effect at this scale as a driver of diatom community patterns (Bottin et al., 2016; Goldenberg Vilar et al., 2014; Jamoneau et al., 2018), they failed to prove it due to the lack of a robust technique to distinguish dispersal limitation and mass effect.

Mass effect is indeed challenging to describe. If a combination of spatial and trait analyses can allow to reveal the importance of mass effect in structuring active disperser community, e.g., benthic macro-invertebrates (Brown and Swan, 2010), this type of approach cannot be used for passive dispersers such as diatoms, which hampers our ability to integrate the potential influence of mass effect on bioassessment. However, Leboucher et al. (*in prep.*) partially rule out this problem for benthic diatoms with the development of an innovative method based on species distribution modeling, with environmental and spatial factors computed from asymmetric eigenvector maps, coupled with negative co-occurrence patterns. This work allow for the detection of species influenced by mass effect and finally provided a list of 40 diatom species to consider more carefully when assigning ecological status to water bodies.

Based on this new framework (Leboucher et al., *in prep.*), we propose here to concretely integrate the influence of mass effect in diatom-based assessment methods. More specifically, we aim to generate an alternative version of the Biological Diatom Index (BDI) (Coste et al., 2009; Lenoir and Coste, 1996), a standardized method (AFNOR, 2007) created and used in France for the monitoring of watercourses quality. We will then compare the distributions of scores and ecological status assessments obtained with the new and original BDI versions, and their responses to environmental conditions.

## **Materials and methods**

### **Diatom and environmental condition datasets**

We analyzed a benthic diatom dataset from France collected between 2007 and 2013, including 9487 samples from 3913 spatially distinct localities (Fig. 5.1). Diatoms were sampled on hard substrates between May and October, in order to reduce substrate and seasonal variabilities in species assemblages, following the standardized protocol (NFT

90-354, AFNOR, 2007). Cells were identified at 1000 x magnification by examining 400 cleaned diatom frustules in permanent slides using among others Krammer and Lange-Bertalot (1986-1991) and Lange-Bertalot (1995-2015, 2000-2013) as identification references. A taxonomic homogenization at the species level was performed with Omnidia 5.3 software (Lecointe et al., 1993).

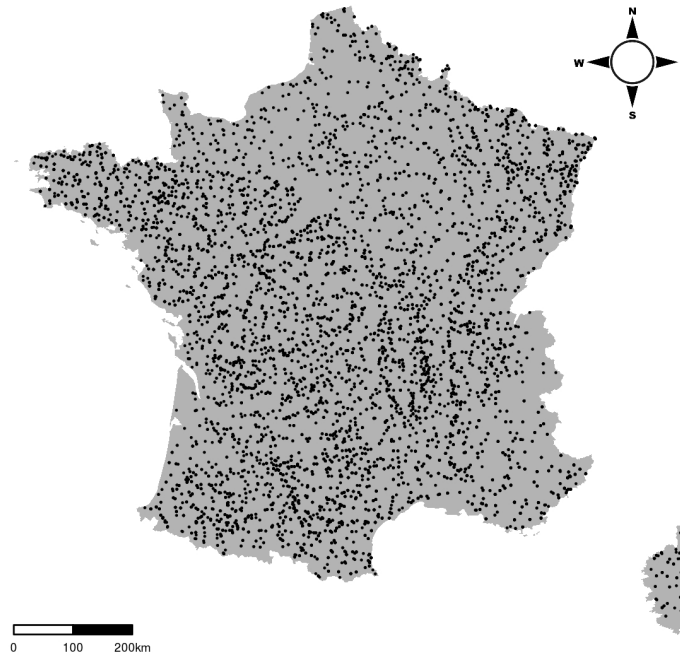


FIGURE 5.1 – Map of the 3913 spatially distinct localities in France.

Corresponding physico-chemical data were gathered by the French Water Agencies. Physico-chemical variables included water pH, specific conductance (at 20 °C,  $\text{mS cm}^{-1}$ ), biological oxygen demand (mg of  $\text{O}_2$  consumed per liter during 5 days of incubation at 20 °C), concentrations of total phosphorous ( $\text{mg L}^{-1}$  of  $^{15}\text{P}$ ), orthophosphate ( $\text{mg L}^{-1}$  of  $^{15}\text{P}$ ), ammonium ( $\text{mg L}^{-1}$  of  $^{15}\text{N}$ ), nitrate ( $\text{mg L}^{-1}$  of  $^{15}\text{N}$ ) and nitrite ( $\text{mg L}^{-1}$  of  $^{15}\text{N}$ ). We used the median values of all environmental variable measurements obtained during the 30 days before and the 15 days after the diatom sampling date.

### Biological Diatom Index (BDI)

$\text{BDI}_{2007}$  (Coste et al., 2009; Lenoir and Coste, 1996) was calculated using the sum of the presence probability of all indicator taxa in each quality class  $j$  (7 predefined classes)

following the equation

$$BDI = \sum_{j=1}^7 j \cdot \frac{\sum_{i=1}^n A_i \cdot V_i \cdot P_{ij}}{\sum_{i=1}^n A_i \cdot V_i}$$

where  $A_i$  is the abundance of taxon  $i$ ,  $V_i$  is the degree of stenoecy of the taxon  $i$ ,  $P_{ij}$  is the presence probability of the taxon  $i$  for the quality class  $j$ .

The value of BDI was related to a reference value for the type of river considered to calculate the Ecological Quality Ratio (EQR) and was then related to the regulatory pre-established class boundaries in order to evaluate the site ecological status (JORF, 2015). This version of BDI was called the original BDI.

A second version of the BDI was calculated after removing the 40 species previously identified as influenced by mass effect (Tab. 5.1, Leboucher et al., *in prep.*) from the list of indicator taxa. This version was called modified BDI.

## Relationships between BDI indices and environmental conditions

We performed a principal component analysis (PCA) on power-transformed physico-chemical variables. We used Box-cox maximum likelihood estimation (Box and Cox, 1964) to decide the best power-transformation for approximating normality. We simplified the analysis by retaining only the components with eigenvalues greater than those randomly generated by a broken-stick model (Jackson, 1993)(Legendre and Legendre, 2012). We used samples coordinates on the principal component axes (PC), which were retained by the broken-stick model, to evaluate the sample site position on the environmental gradient of the dataset.

We used the Akaike information criterion (minimum theoretical information criterion, AIC) (Akaike, 1974) to compare linear models between both versions of the EQR scores (BDI) and the PCA coordinates. Smallest value of AIC indicates the best model, and therefore, the BDI version with the best response towards environmental conditions.

| ID   | Species                                                                      |
|------|------------------------------------------------------------------------------|
| DVUL | <i>Diatoma vulgaris</i> Bory                                                 |
| ADEU | <i>Achnanthyidium eutrophilum</i> (Lange-Bertalot) Lange-Bertalot            |
| ACOP | <i>Amphora copulata</i> (Kützing) Schoeman & Archibald                       |
| AOVA | <i>Amphora ovalis</i> (Kützing) Kützing                                      |
| APED | <i>Amphora pediculus</i> (Kützing) Grunow                                    |
| CNTH | <i>Cocconeis neothumensis</i> Krammer                                        |
| CPED | <i>Cocconeis pediculus</i> Ehrenberg                                         |
| CPLA | <i>Cocconeis placentula</i> Ehrenberg                                        |
| COPL | <i>Cocconeis pseudolineata</i> (Geitler) Lange-Bertalot                      |
| DMON | <i>Diatoma moniliformis</i> Kützing                                          |
| DOCU | <i>Diploneis oculata</i> (Brebisson) Cleve                                   |
| ENLB | <i>Encyonema lange-bertalotii</i> Krammer                                    |
| ECPM | <i>Encyonopsis minuta</i> Krammer & Reichardt                                |
| EOMI | <i>Eolimna minima</i> (Grunow) Lange-Bertalot                                |
| FPYG | <i>Fallacia pygmaea</i> (Kützing) Stickle & Mann                             |
| FSLU | <i>Fallacia sublucidula</i> (Hustedt) Mann                                   |
| GMIN | <i>Gomphonema minutum</i> (Agardh)                                           |
| GOLI | <i>Gomphonema olivaceum</i> (Hornemann) Brubisson                            |
| GPUM | <i>Gomphonema pumilum</i> Reichardt & Lange-Bertalot                         |
| GYAT | <i>Gyrosigma attenuatum</i> (Kützing) Rabenhorst                             |
| KAPL | <i>Karayevia ploenensis</i> (Hustedt) Bukhtiyarova                           |
| LGOE | <i>Luticola goeppertiana</i> (Bleisch in Rabenhorst) Mann                    |
| NANT | <i>Navicula antonii</i> Lange-Bertalot                                       |
| NCPR | <i>Navicula capitatoradiata</i> Germain                                      |
| NCTE | <i>Navicula cryptotenella</i> Lange-Bertalot                                 |
| NGER | <i>Navicula germainii</i> Wallace                                            |
| NAMP | <i>Nitzschia amphibia</i> Grunow                                             |
| NCPL | <i>Nitzschia capitellata</i> Hustedt in Schmidt                              |
| NFON | <i>Nitzschia fonticola</i> Grunow                                            |
| NREC | <i>Nitzschia recta</i> Hantzsch in Rabenhorst                                |
| NSOC | <i>Nitzschia sociabilis</i> Hustedt                                          |
| NSUA | <i>Nitzschia subacicularis</i> Hustedt in Schmidt                            |
| NTUB | <i>Nitzschia tubicola</i> Grunow                                             |
| PGRN | <i>Planothidium granum</i> (Hohn & Hellerman) Lange-Bertalot                 |
| PRST | <i>Planothidium rostratum</i> (Oestrup) Lange-Bertalot                       |
| PTCO | <i>Platessa conspicua</i> (Mayer) Lange-Bertalot                             |
| PBTG | <i>Pseudostaurosira brevistriata</i> (Grunow in Van Heurck) Williams & Round |
| RUNI | <i>Reimeria uniseriata</i> Sala Guerrero & Ferrario                          |
| RABB | <i>Rhoicosphenia abbreviata</i> (Agardh) Lange-Bertalot                      |
| SIDE | <i>Simonsenia delognei</i> Lange-Bertalot                                    |

TABLE 5.1 – list of the 40 key taxa identified as influenced by mass effect from Leboucher et al. (*in prep.*).

## Comparison of ecological status assessment

We compared the distribution of ecological status assessment between the original and the modified versions of BDI using exact binomial tests. We also analyzed the changes in ecological status assessment from one version to another, and whether these changes were correlated with the nutrient gradient (PC1 coordinates).

All data analyses were performed with R 3.4.4 (R Core Team, 2018), using the ‘vegan’ package (Oksanen et al., 2018).

## Results

Following the broken-stick model, only the first principal component (PC1) axis was significant explaining 40 % of the total variance (Fig. 5.2). PC1 represented a gradient of nutrient enrichment mainly driven by phosphorous (orthophosphates and total phosphorous), nitrogen (nitrites and ammonia) and the biological oxygen demand. This is consistent with the expected response of BDI towards environmental conditions (Coste et al., 2009). Only samples coordinates on PC1 were then used to compare the response of the the two BDI indices to the environmental gradient. AIC of linear models between BDI indices and the nutrient gradient (PC1 coordinates) was smallest for the modified version (AIC =  $-51241$ , Adj. R<sup>2</sup> = 0.28, p-value < 0.001) compared to the original one (AIC =  $-50967$ , Adj. R<sup>2</sup> = 0.25, p-value < 0.001) indicating a better relationship between the modified version scores and the environmental conditions.

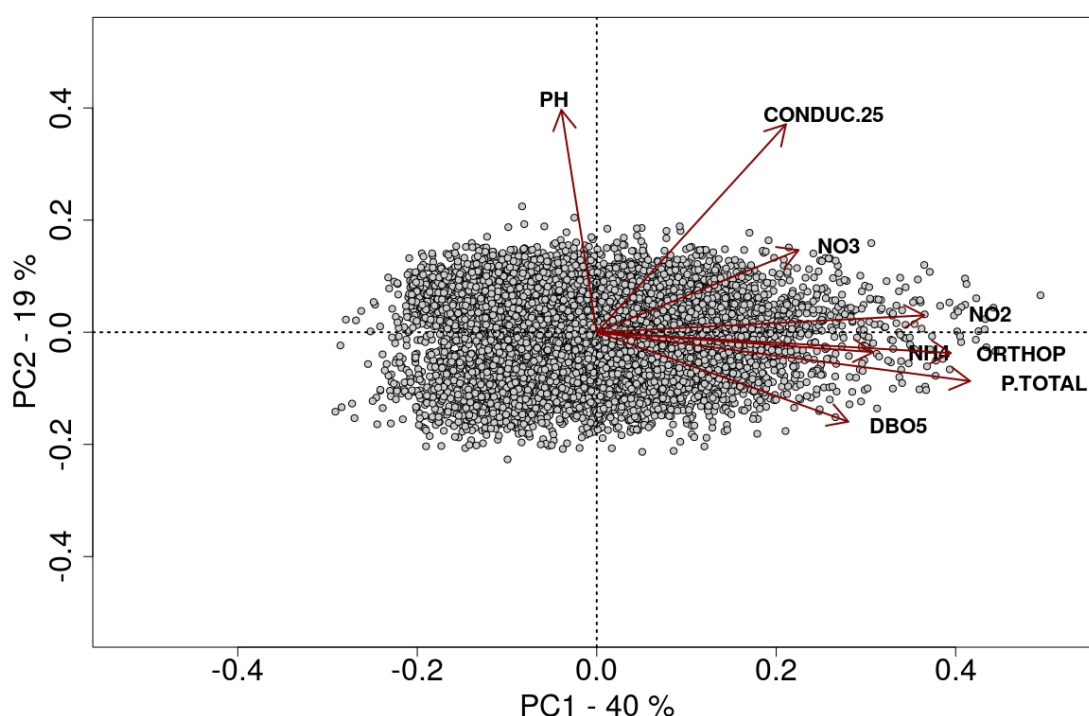


FIGURE 5.2 – Principal component analysis biplot for axes 1 and 2 using pH, specific conductance (CONDUCT.25), biological oxygen demand (DBO5), nitrite (NO<sub>2</sub>), ammonia (NH<sub>4</sub>), total phosphorous (P.TOTAL) and orthophosphate (ORTHOP) concentrations. Environmental variables are shown with arrows.

Ecological status showed significant differences (exact binomial tests,  $p\text{-value} < 0.001$ ) between the original and the modified BDI versions (Fig. 5.3). The modified BDI version presented a more homogeneous distribution of samples between status classes compared to the original version. The number of sites belonging to average status (Good and Moderate) decreased in favor of more extreme status (High, Poor and Bad).

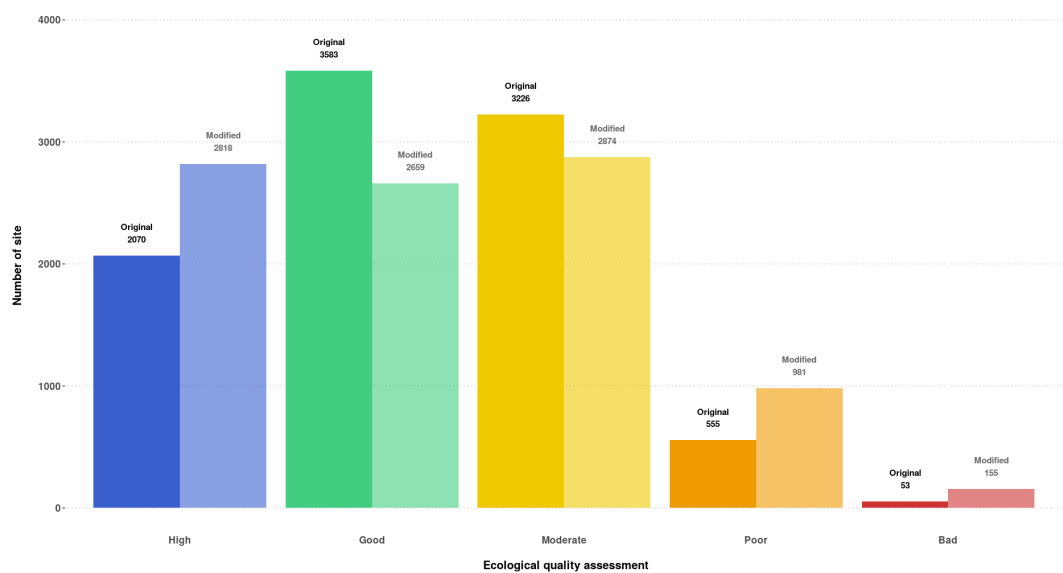


FIGURE 5.3 – Distribution of the ecological status assessment for the original and modified versions of BDI.

If the samples originally belonging to High and Bad quality status did not change with the modified version, samples belonging to good quality status were in majority spread over High, Good and Moderate classes with the modified version ; samples originally belonging to Moderate quality status were spread over Good, Moderate and Poor classes ; and samples originally belonging to Poor quality status were spread over Poor and Bad classes (Tab. 5.2).

|              |          | Modified BDI |      |          |      |     |
|--------------|----------|--------------|------|----------|------|-----|
| Original BDI | High     | High         | Good | Moderate | Poor | Bad |
|              | Good     | 2061         | 9    | 0        | 0    | 0   |
|              | Moderate | 745          | 2353 | 483      | 2    | 0   |
|              | Poor     | 12           | 297  | 2379     | 536  | 2   |
|              | Bad      | 0            | 0    | 12       | 443  | 100 |
|              |          | 0            | 0    | 0        | 0    | 53  |

TABLE 5.2 – Classification among ecological status assessment for the original and the modified versions of BDI (number of samples)

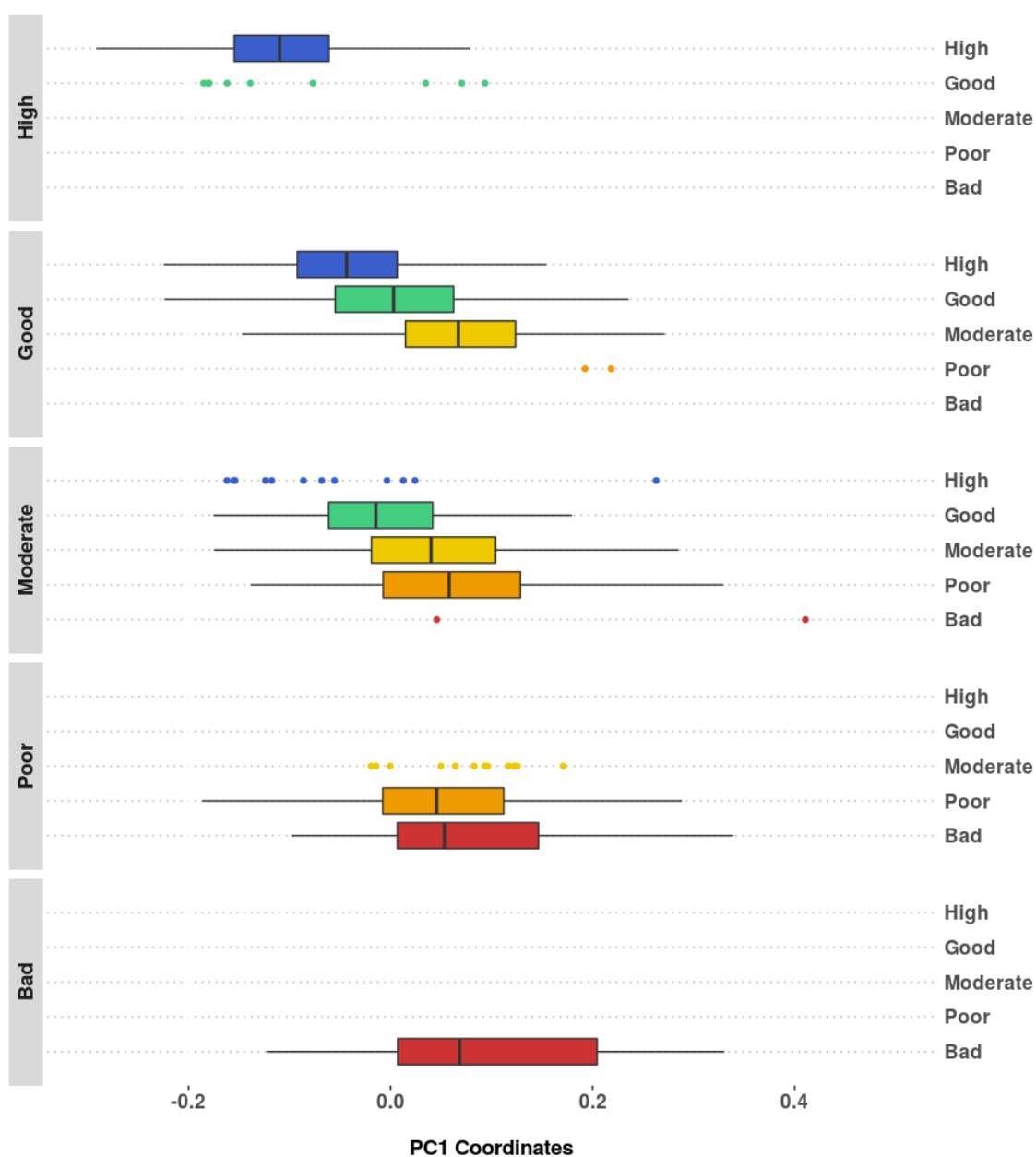


FIGURE 5.4 – Boxplot of the coordinates on PC1 in function of the ecological status assessment of the original BDI version (grey boxes on the left) and in function of the ecological status assessment of the modified BDI version (on the right).

These significant differences between the original and the modified version allowed for a better discrimination of ecological status assessments (Fig. 5.4). Indeed, the dispersion of samples splitted from the original class into the different classes obtained with the modified version was correlated with the nutrient gradient observed on PCA. Samples upgraded with the new BDI version showed lower nutrient concentrations (low PC1 scores) whereas

samples downgraded showed higher nutrient concentrations (high PC1 scores).

## Discussion

Removing species influenced by mass effect from the list of the indicator taxa led to a new version of the BDI which improved the relationships between environmental conditions and index scores, and to a better discrimination of samples ecological status. If previous literature had considered that mass effect could mask the effect of the local environment resulting in a potential improper evaluation of water quality (Heino et al., 2015; Jamoneau et al., 2018; Leibold et al., 2004), our study is actually the first, to our knowledge, to demonstrate this. We discuss the implications of these findings below.

Spatial structuring of community has long been underestimated by ecologists and its importance has only recently been stressed by the metacommunity theory (Leibold et al., 2004) and most of recent studies integrated the measurement of the 'pure' spatial component. Although this process has been evoked as a driver of community patterns by several studies, including diatom communities, only the development of a novel technique by Leboucher et al. (*in prep.*) allowed for the detection of species influenced by mass effect, which had never been done before. Among the studies that have mentioned the mass effect, Bottin et al. (2016) found for example that co-occurrence patterns departed from both random and environmentally predicted structure, and related these overall aggregative patterns potentially to mass effect. Using beta-diversity measurements, Goldenberg Vilar et al. (2014) also reported that under mesotrophic conditions, purely spatial processes play a prominent role in structuring diatom communities. As microorganisms are unlikely to be dispersal-limited at the scale of their study (less than 6 km), they interpreted this pattern as a result of mass effect. They also noticed a related increase of the relative abundance of motile species in such conditions (Passy, 2007). Such a result is consistent with Jamoneau et al. (2018) who examined the variation in total beta-diversity and its components (turnover and nestedness) and suggested that the influence of spatial factors mid- and down-stream, together with a decrease of beta- and gamma-diversity, and an increase of nestedness, was due to mass effect, with here also motile species being favored by this process at the expense of low profile species. Surprisingly, Leboucher et al. (*in prep.*) did not notice that a particular ecological guild was prone to be affected by mass effect, which is not in line with Goldenberg Vilar et al. (2014) and Jamoneau et al. (2018) concerning motile species.

This difference in results for ecological guilds may be due to a confusion between species influenced by the mass effect and generalists. Gutiérrez-Cánovas et al. (2013) related eutrophication to the elimination of sensitive species (specialist) and to the subsequent replacement by a few generalist species usually sharing similar traits. Such patterns were also described by Leboucher et al. (*in prep.*) in diatom communities who revealed that

beta-diversity was dependent on nutrient level and that greater enrichment more strongly depressed beta-diversity. This pattern was attributable to the fact that among impacted streams, relatively fewer new species (particularly those with sensitive/specialist traits) were added to the regional species pool compared to unimpacted streams. Acute and persistent anthropogenic stress, such as eutrophication, may offer insufficient time for adaptation by specialists and potentially eliminate them from the species pool (Gutiérrez-Cánovas et al., 2013; Kaspari et al., 2010). The loss of specialists benefits subsequent colonization by more physiologically plastic generalists, which are normally poor competitors for resources and niche space compared to specialists (Bolnick et al., 2007; Olden and Rooney, 2006). However, Leboucher et al. (*in prep.*) also reported a 'pure' spatial structuring of some diatoms communities in nutrient impacted conditions. This observation was related to dispersal processes, mass effect or dispersal limitation, but without untangling these two processes because they both can be inferred from correlations between spatial variables and community composition. By taking into account niche breadths in their study, the methodological development provided by Leboucher et al. (*in prep.*) makes it possible to better understand the specific influence of the mass effect on communities and the Goldenberg Vilar et al. (2014) and Jamoneau et al. (2018) observations may then be the result of the combined effect of the replacement of specialists by generalists, with in parallel the emergence of species influenced by the mass effect.

Leboucher et al. (*in prep.*) results tented anyway to conclude that mesotrophic conditions were favorable to mass effect. In line with this, the majority of the species detected as influenced by this process are traditionally considered as mesosaprobic and/or mesotrophic, typical of good to medium water quality (Van Dam et al., 1994), but not nutrient limited. Such result seems totally consistent if we consider that species dispersal is in the mass effect theory associated to high population densities. Diatom capacity to produce high population densities indeed has already been related to nutrient levels as low level of phosphorus (Bothwell, 1989; Chételat et al., 1999) and nitrogen (Stevenson et al., 2006) limiting diatom community growth.

Our results largely supported the idea that mass effect induce a potential improper evaluation of water quality. The comparison of the two versions of the BDI shown in every respect that taking into account the mass effect leads to a more acute ecological assessment of river status. The modified version of the BDI strengthened the correlation between index scores and the local environmental conditions and better discriminated samples ecological status. This advocates for an application of our method to any diatom-based indices, and by extension to any biotic index based on active or passive dispersers.

## References

- AFNOR, 2007. Norme NFT 90-354 - Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD).
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723.
- Bolnick, D.I., Svanbäck, R., Araújo, M.S., Persson, L., 2007. Comparative support for the niche variation hypothesis that more generalized populations also are more heterogeneous. *Proceedings of the National Academy of Sciences* 104, 10075–10079.
- Bothwell, M.L., 1989. Phosphorus-Limited Growth Dynamics of Lotic Periphytic Diatom Communities : Areal Biomass and Cellular Growth Rate Responses. *Canadian Journal of Fisheries and Aquatic Sciences* 46, 1293–1301.
- Bottin, M., Soininen, J., Alard, D., Rosebery, J., 2016. Diatom Cooccurrence Shows Less Segregation than Predicted from Niche Modeling. *PLOS ONE* 11, e0154581.
- Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26, 211–252.
- Brown, B.L., Swan, C.M., 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology* 79, 571–580.
- Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O., Patrick, C., 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society* 30, 310–327.
- Chételat, J., Pick, F.R., Morin, A., Hamilton, P.B., 1999. Periphyton biomass and community composition in rivers of different nutrient status. *Canadian Journal of Fisheries and Aquatic Sciences* 56, 560–569.
- Coste, M., Boutry, S., Tison-Rosebery, J., Delmas, F., 2009. Improvements of the Biological Diatom Index (BDI) : Description and efficiency of the new version (BDI-2006). *Ecological Indicators* 9, 621–650.

- Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* 8, 1175–1182.
- Dong, X., Li, B., He, F., Gu, Y., Sun, M., Zhang, H., Tan, L., Xiao, W., Liu, S., Cai, Q., 2016. Flow directionality, mountain barriers and functional traits determine diatom metacommunity structuring of high mountain streams. *Scientific Reports* 6.
- Goldenberg Vilar, A., van Dam, H., van Loon, E.E., Vonk, J.A., van Der Geest, H.G., Admiraal, W., 2014. Eutrophication decreases distance decay of similarity in diatom communities. *Freshwater Biology* 59, 1522–1531.
- Gutiérrez-Cánovas, C., Millán, A., Velasco, J., Vaughan, I.P., Ormerod, S.J., 2013. Contrasting effects of natural and anthropogenic stressors on beta diversity in river organisms. *Global Ecology and Biogeography* 22, 796–805.
- Heino, J., Grönroos, M., Ilmonen, J., Karhu, T., Niva, M., Paasivirta, L., 2013. Environmental heterogeneity and beta diversity of stream macroinvertebrate communities at intermediate spatial scales. *Freshwater Science* 32, 142–154.
- 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. *Freshwater Biology* 60, 845–869.
- Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis : A Comparison of Heuristical and Statistical Approaches. *Ecology* 74, 2204–2214.
- Jamoneau, A., Passy, S.I., Soininen, J., Leboucher, T., Tison-Rosebery, J., 2018. Beta diversity of diatom species and ecological guilds : Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology* 63, 62–73.
- Johnson, R.K., Furse, M.T., Hering, D., Sandin, L., 2007. Ecological relationships between stream communities and spatial scale : implications for designing catchment-level monitoring programmes. *Freshwater Biology* 52, 939–958.
- JORF, 2015. Arrêté du 27 juillet 2015 modifiant l'arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement. *Journal Officiel de la République Française*.
- Kaspari, M., Stevenson, B.S., Shik, J., Kerekes, J.F., 2010. Scaling community structure : how bacteria, fungi, and ant taxocenes differentiate along a tropical forest floor. *Ecology* 91, 2221–2226.
- Legendre, P., Legendre, L.F.J., 2012. *Numerical Ecology*. Elsevier.

- 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., Gonzalez, A., 2004. The metacommunity concept : a framework for multi-scale community ecology. *Ecology Letters* 7, 601–613.
- Lenoir, A., Coste, M., 1996. Development of a practical diatom index of overall water quality applicable to the French national water board network. Use of algae for monitoring rivers , 29–43.
- Mouquet, N., Loreau, M., 2003. Community Patterns in Source Sink Metacommunities. *The American Naturalist* 162, 544–557.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. *vegan* : Community ecology package. R package version , 117–118.
- Olden, J.D., Rooney, T.P., 2006. On defining and quantifying biotic homogenization. *Global Ecology and Biogeography* 15, 113–120.
- Passy, S.I., 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquatic Botany* 86, 171–178.
- Passy, S.I., 2009. The relationship between local and regional diatom richness is mediated by the local and regional environment. *Global Ecology and Biogeography* 18, 383–391.
- Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., Schmutz, S., 2006. Assessing river biotic condition at a continental scale : a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology* 43, 70–80.
- Potapova, M.G., Charles, D.F., 2002. Benthic diatoms in USA rivers : distributions along spatial and environmental gradients. *Journal of Biogeography* 29, 167–187.
- Prygiel, J., Coste, M., Bukowska, J., 1999. Review of the major diatom-based techniques for the quality assessment of rivers-State of the art in Europe. Use of algae for monitoring rivers 3, 224–238.
- R Core Team, 2018. *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Shmida, A., Wilson, M.V., 1985. Biological Determinants of Species Diversity. *Journal of Biogeography* 12, 1–20.
- Smucker, N.J., Vis, M.L., 2011a. Diatom biomonitoring of streams : Reliability of reference

sites and the response of metrics to environmental variations across temporal scales. *Ecological Indicators* 11, 1647–1657.

Smucker, N.J., Vis, M.L., 2011b. Spatial factors contribute to benthic diatom structure in streams across spatial scales : Considerations for biomonitoring. *Ecological Indicators* 11, 1191–1203.

Soininen, J., Jamoneau, A., Rosebery, J., Passy, S.I., 2016. Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography* 25, 940–950.

Stevenson, R.J., Rier, S.T., Riseng, C.M., Schultz, R.E., Wiley, M.J., 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria, in : *Advances in Algal Biology : A Commemoration of the Work of Rex Lowe*. Springer, Dordrecht. *Developments in Hydrobiology*, pp. 149–165.

Tison, J., Park, Y.S., Coste, M., Wasson, J.G., Ector, L., Rimet, F., Delmas, F., 2005. Typology of diatom communities and the influence of hydro-ecoregions : A study on the French hydrosystem scale. *Water Research* 39, 3177–3188.

Tonkin, J.D., Altermatt, F., Finn, D.S., Heino, J., Olden, J.D., Pauls, S.U., Lytle, D.A., 2018. The role of dispersal in river network metacommunities : Patterns, processes, and pathways. *Freshwater Biology* 63, 141–163.

Van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Netherlands Journal of Aquatic Ecology* 28, 117–133.

Vanormelingen, P., Verleyen, E., Vyverman, W., 2008. The diversity and distribution of diatoms : from cosmopolitanism to narrow endemism. *Biodiversity and Conservation* 17, 393–405.

Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Vijver, B.V., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P., Wever, A.D., 2007. Historical Processes Constrain Patterns in Global Diatom Diversity. *Ecology* 88, 1924–1931.

## **Acknowledgements**

We thank all French Water Agencies for data sharing. We gratefully acknowledge financial support from AFB (French Agency for Biodiversity).



## 6 Apports, limites et perspectives du travail réalisé

### 6.1 Bilan général

#### 6.1.1 Influences des processus déterministes et stochastiques sur la structuration des communautés de diatomées benthiques

L'utilisation d'une base de données intercontinentale ainsi que la mise en place d'un cadre multi-échelles explicite nous ont permis de mieux appréhender les facteurs influençant la diversité  $\beta$  au sein des communautés diatomiques. Il en est ressorti que la diversité  $\beta$  augmentait avec l'échelle spatiale conformément à ce qui avait pu être rapporté dans la littérature (Gabriel et al., 2006; Martiny et al., 2011). Le niveau d'anthropisation (ici, teneur en nutriments) joue cependant un rôle majeur puisque même une légère augmentation du niveau trophique entraîne une forte baisse de la diversité  $\beta$ . Il est à noter ici que la méthode de calcul de l'indice de diversité  $\beta$  n'influe que peu sur les résultats, et que les différents indices également testés donnaient des résultats très similaires. Seule la décomposition de la diversité  $\beta$  en ces deux composants, le *turnover* et la *nestedness* (d'après Baselga, 2010), n'a pas été très concluante mais ce résultat était attendu (Jamoneau et al., 2018; Soininen et al., 2018). De par la structure des communautés de diatomées et le nombre d'espèces décrites, la part de *nestedness* reste toujours très faible et les résultats qui en découlent relativement peu fiables. Notre hypothèse initiale était que le niveau d'anthropisation, en appauvrissant le *pool* d'espèces et en favorisant l'émergence d'espèces généralistes, favoriserait de forte valeurs de *nestedness*. Bien que des différences significatives dans ce sens aient été observées lors de nos essais, celles-ci variaient très peu (de 0.033 dans les conditions non impactées par les nutriments, à 0.037 dans les conditions les plus impactées) et représentaient moins de 5% de la diversité  $\beta$  totale, et il était difficile de conclure entre réalité écologique ou simple artefact mathématique.

La théorie derrière ces observations semble toutefois concordante avec nos résultats puisque nous avons pu relier la chute de la diversité  $\beta$  dans les conditions impactées à un remplacement d'espèces spécialistes par des espèces plus généralistes. Ce rem-

placement peut s'expliquer par les différences existant entre les gradients de stress naturels et d'origine humaine. Si les gradients naturels de stress (gradient de salinité par exemple) existent depuis des milliers d'années permettant la spécialisation progressive des espèces à leur environnement local, ce n'est pas le cas des gradients de stress d'origine humaine dont la persistance dans le temps n'est certainement pas encore assez importante pour permettre l'émergence d'espèces particulièrement adaptées à ces milieux, éliminant de fait les espèces spécialistes présentes avant l'émergence du stress (Gutiérrez-Cánovas et al., 2013; Kaspari et al., 2010). Cette disparition des espèces spécialistes profiterait ainsi à quelques généralistes possédant une plasticité phénotypique plus importante et qui sont habituellement de mauvais compétiteurs (Bolnick et al., 2007; Olden and Rooney, 2006).

L'utilisation d'un modèle nul nous a permis de relier ce résultat à l'influence relative de la diversité  $\gamma$  et des processus locaux structurant les communautés. Bien que ces derniers dominent à toutes les échelles d'observation, l'appauvrissement des *pools* d'espèces et la dominance des espèces généralistes dans les milieux fortement impactés par l'anthropisation semble augmenter l'importance de la diversité  $\gamma$  dans la structuration des communautés, indiquant que les communautés dans les milieux impactés résultent également d'un échantillonnage aléatoire d'espèces généralistes présentes dans le *pool* régional (Karp et al., 2012; Kraft et al., 2011). Il est cependant délicat de véritablement parler de *pool* régional lorsque l'on évoque la diversité  $\gamma$ . Celle-ci n'est en effet souvent qu'un sous échantillon contraint par le nombre de sites pris en considération du véritable *pool* d'espèces qui est difficile à appréhender.

Il est également apparu que l'influence du *species sorting* diminuait dans les milieux impactés, au profit de processus purement spatiaux. Il n'a malheureusement pas été possible de conclure entre effet de masse et dispersion limitée, puisque ces deux processus peuvent générer des signaux spatiaux. Si la baisse de la diversité  $\beta$  dans les milieux impactés peut laisser suggérer qu'il puisse s'agir de l'effet de masse, cette hypothèse ne pouvait pas être vérifiée. Le développement méthodologique présenté dans le chapitre 4 qui nous a permis de mesurer l'influence de l'effet de masse sur la distribution des espèces de diatomées, aurait pu nous permettre d'aller plus loin sur cette question et de conforter, ou non, cette hypothèse. Une solution aurait pu être de regarder si les espèces détectées comme étant fortement influencées par l'effet de masse avec cette nouvelle méthodologie, étaient les espèces qui contribuaient le plus à la diversité  $\beta$  (Heino and Grönroos, 2017) dans les sites impactés.

### 6.1.2 Développement méthodologique pour la mise en évidence de l'influence de l'effet de masse sur la distribution des espèces

Le développement de cette nouvelle méthodologie est un des points marquants de cette thèse. Elle permet, pour la première fois à notre connaissance, de détecter l'effet de ce processus sur la distribution des espèces de diatomées. Si l'effet de masse, en particulier dans les communautés diatomiques, avait déjà été évoqué par quelques études, les preuves avancées étaient encore indirectes et non mesurées à l'échelle spécifique. Parmi ces études, l'observation d'une baisse de la diversité  $\beta$  couplée à des signaux purement spatiaux a été plusieurs fois utilisée pour évoquer l'influence de l'effet de masse sur les communautés (Goldenberg Vilar et al., 2014; Jamoneau et al., 2018). Goldenberg Vilar et al. (2014) ont notamment trouvé que l'influence des processus spatiaux sur la structuration des communautés de diatomées était particulièrement importante dans les conditions mésotrophiques et corrélée à une augmentation de la proportion des espèces *motile* dans les communautés (Passy, 2007). Comme les micro-organismes ne sont pas connus pour subir de la dispersion limitée à petite échelle (moins de 6 km), l'effet de masse avait été évoqué pour expliquer cette observation. Ces résultats sont cohérents avec ceux obtenus par Jamoneau et al. (2018). En examinant les variations des composants de la diversité  $\beta$  (*turnover* et *nestedness*), ces auteurs ont mis en évidence que de forts signaux spatiaux dans les sections en aval des cours d'eaux étudiés, étaient associés à une baisse du *turnover* et une augmentation de la *nestedness* suggérant ainsi une forte influence de l'effet de masse sur la structuration des communautés dans ces conditions. Comme dans l'étude de Goldenberg Vilar et al. (2014), une augmentation de la proportion des espèces *motile* a également été observée, au détriment des espèces *low profile*.

Si la détection de l'effet de masse dans des conditions méso- à eutrophe semble cohérente avec nos résultats, il est étonnant de ne pas retrouver d'effet sur les guildes écologiques de Passy (2007) dans notre étude. Une explication possible de ce résultat serait une confusion avec l'augmentation des espèces généralistes dans les conditions méso- à eutrophes. En effet, dans les résultats du chapitre 3, nous avons mis en évidence qu'une augmentation de la teneur en nutriments induisait un remplacement des espèces spécialistes par des espèces généralistes. Il est donc tout à fait envisageable que les observations de Goldenberg Vilar et al. (2014) et de Jamoneau et al. (2018) concernant les guildes diatomiques soient à la fois la conséquence de l'influence de l'effet de masse et de ce remplacement des espèces spécialistes, expliquant ainsi pourquoi nous n'avons pas retrouvé de tels patrons dans notre étude.

Quoi qu'il en soit, ces résultats montrent qu'une meilleure intégration des processus spatiaux est aujourd'hui nécessaire en bio-indication. Quasiment toutes les espèces dont la distribution est influencée par l'effet de masse rentrent actuellement dans le calcul de plusieurs indices basés sur les diatomées. Ces espèces devraient être prises en compte avec plus de précautions dans les indices, soit en leur attribuant un poids moins important dans le calcul, soit en les retirant tout simplement de la liste des taxons indicateurs. La

présence de ces espèces peut en effet masquer l'effet de l'environnement local (Heino et al., 2015; Leibold et al., 2004) et ainsi aboutir à une évaluation erronée de la qualité de l'eau. D'une manière plus générale, ce type de démarche devrait être élargie à l'ensemble des organismes servant aujourd'hui de bio-indicateurs puisque l'effet de masse n'est pas un processus propre aux diatomées (Brown and Swan, 2010).

L'une des limites à la généralisation de cette approche est toutefois l'utilisation des BRT pour mesurer l'influence de l'effet de masse sur la distribution d'une espèce. Ce type d'analyse exige en effet une quantité de données relativement importante ainsi qu'une occurrence relativement élevée de l'espèce car ces modèles sont peu adaptés aux espèces rares (Brotons et al., 2004).

### **6.1.3 Influence de l'effet de masse en bio-indication : exemple de l'Indice Biologique Diatomée**

Pour répondre à ce dernier objectif de la thèse, nous avons choisi de retirer de la liste des taxons indicateurs de l'IBD les 40 espèces dont la distribution était fortement influencée par l'effet de masse. Cette version modifiée de l'IBD permet une réponse plus fine de l'indice à l'environnement local et ainsi une meilleure évaluation écologique.

Une autre solution aurait pu être de donner à ces espèces un poids plus faible, autrement dit faire baisser leur valence écologique, dans le calcul. Nous n'avons pas retenu cette option car les valeurs de valence écologique ont été obtenues grâce à des analyses portant sur l'ensemble des espèces. Changer arbitrairement ces valeurs n'avait donc aucun sens écologiquement parlant.

Parmi ces espèces, certaines avaient pu être classées dans différents groupes en fonction des bassins versants étudiés. Ce résultat mériterait d'être approfondi même s'il n'est pas surprenant en soit. L'effet de masse dépend en effet des conditions locales existant dans chaque bassin et il se peut que certains bassins ne regroupent, par exemple, que des environnements favorables pour une espèce donnée, auquel cas nous ne sommes pas censés retrouver de structuration purement spatiale de l'espèce étudiée.

La principale différence entre les deux versions de l'IBD repose essentiellement sur une meilleure discrimination des notes dans les états intermédiaires (mauvais, moyen et bon états). Ce résultat semble cohérent avec ceux du chapitre précédent puisque ces classes d'états correspondent globalement aux milieux mésotrophes à eutrophes. Le flou induit par l'effet de masse est loin d'être négligeable dans ces milieux puis qu'environ 25 % des échantillons diatomiques ont changé de classe d'état entre la version originale de l'IBD et cette nouvelle version.

Ces résultats ne font que souligner l'importance de la prise en compte des processus de dispersion dans l'évaluation de l'état écologique et dans le développement de nouveaux indices de bio-indication.

## **6.2 Perspectives de recherche**

### **6.2.1 Base de données**

Bien que très conséquentes, les bases de données françaises et américaines utilisées lors de cette étude gagneraient à être complétées par de nouvelles données issues d'autres pays. Lors de l'utilisation des fenêtres paysagères dans le chapitre 3, une des limites pour l'exploitation de nos résultats était due au décalage entre les échelles des fenêtres en France (entre 200 km et 800 km) et aux États-Unis (entre 800 km et 2000 km). Ce manque de recouvrement entre les deux pays ne nous a pas permis de conclure quant aux processus sous-jacent aux différences de patrons de diversité  $\beta$  observés. Si aux États-Unis, la densité de données était insuffisante pour nous permettre de créer des fenêtres à plus petite échelle, en France, il n'a pas été possible d'augmenter la taille des fenêtres au delà de 800 km à cause de la taille du territoire métropolitain. L'utilisation de données issues des pays limitrophes, comme la Belgique, l'Espagne, la Suisse ou encore l'Italie, aurait été un atout considérable pour cette étude.

### **6.2.2 Occupation des sols et modèles hydrologiques**

L'écologie des communautés de micro-organismes a longtemps souffert de n'être considérée que sous le prisme de la sensibilité ou de la tolérance des espèces à leur environnement local, bien souvent limité à des mesures physico-chimiques ponctuelles. Cette ponctualité des mesures physico-chimiques est un véritable problème lorsque l'on cherche à étudier l'impact de pollutions diffuses (eutrophisation notamment) dans les cours d'eau puisque ce type de prélèvement n'intègre pas la variabilité temporelle des paramètres mesurés (Bernard et al., 2019). Si l'échantillonnage passif permet de mieux intégrer cette variabilité, il demande des conditions d'utilisation particulières, comme la présence de courant, et il serait intéressant de complètement s'affranchir de la mesure des paramètres physico-chimiques.

## Occupation des sols

Une approche possible afin de s'affranchir de ce type de mesures peut être de caractériser l'occupation du sol et de mieux recenser les flux de pollution à l'aide de modèles hydrologiques plus fins (voir par exemple Tudesque et al., 2014). Les pressions appliquées à une station d'échantillonnage diatomique (ou n'importe quel maillon biologique) peuvent alors être déduites de l'occupation des sols, exprimée en pourcentage de surface du territoire voué à telle ou telle activité humaine (forêt, urbanisation, agriculture, etc), et mesurée sur tout ou partie du bassin versant de la station considérée. L'existence de larges bases de données, comme CORINE Land Cover (Union Européenne, 2006), regroupant ce type d'information est une facilitation supplémentaire pour mener à bien ce type d'approche, notamment à très large échelle puisque cet inventaire de l'occupation des terres fournit une information pour 29 pays européens.

Cette mesure de l'occupation du sol aurait pu nous permettre de compléter la liste des prédicteurs environnementaux et ainsi diminuer le risque de mesurer des gradients spatiaux uniquement dus à des gradients environnementaux spatialisés non mesurés. De plus, l'occupation du sol aurait également pu nous fournir une évaluation intégrée sur d'autres types de pressions, comme la présence potentielle d'herbicides dans les cours d'eaux des bassins versants agricoles. Enfin, l'occupation du sol aurait pu nous permettre de mesurer la fragmentation des habitats qui est considérée comme l'une des principales menaces, au même titre que l'eutrophisation par exemple, qui pèsent sur la biodiversité mondiale (Sala et al., 2000).

## Modèles hydrologiques

Concernant les modèles hydrologiques, il serait intéressant d'intégrer les conditions hydrologiques locales (vitesse de courant et présence d'obstacles par exemple) aux modèles, afin d'essayer d'améliorer leurs capacités prédictives. Si les Asymmetric Eigenvector Maps permettent de prendre en compte certaines spécificités des réseaux dendritiques, telles que nous les avons utilisées, elles ne permettent pas de prendre en compte les conditions hydrologiques entre les stations. Ainsi, dans nos analyses, la connectivité mesurée entre deux sites ne tenait pas compte de la distance qui les séparait, ni des conditions hydrologiques locales. Les AEMs permettent pourtant d'intégrer ce type de considérations en donnant un poids différent à chaque "pont" (voir la figure 2.4, tirée de Blanchet et al., 2008a) entre deux sites. Ce type de données est toutefois difficile à obtenir car il nécessite soit un important travail de terrain pour mesurer la vitesse du courant en de multiples endroits (impossible à très large échelle), soit un important travail de géomatique pour calculer les pentes de chaque rivière à partir de modèles numériques de terrain, et de répertorier de manière fine la présence d'obstacles sur les cours d'eau. Cet effort semble pourtant profitable puisqu'il offrirait de nouvelles perspectives de recherches afin de mieux

comprendre la structuration des communautés de diatomées.

### 6.2.3 Structuration des communautés

#### Adaptation intra- et inter-spécifique

Contrairement aux végétaux supérieurs, l'adaptation des communautés de diatomées à leur milieu n'a jamais, à notre connaissance, été appréhendée sous le prisme de la théorie des compromis évolutifs (*trade-off*). Ce type d'approche est évidemment difficile à mettre en place du simple fait de la taille des organismes étudiés et également à cause de l'absence de traits de vie mesurables à l'échelle d'un individu. Ce type d'approche est pourtant l'un des facteurs déterminants dans bien des études d'écologie des communautés, chez les végétaux supérieurs notamment, qui ont montré que les variations phénotypiques à l'échelle intra-spécifique contribuaient fortement à la réponse des communautés végétales aux contraintes environnementales (voir par exemple Jung et al., 2014, 2010; Violle et al., 2012; Whitlock et al., 2007). Lorsque l'on considère le nombre de complexes d'espèces cryptiques chez les diatomées, il apparaît tout à fait probable qu'une adaptation morphologique intra- et inter-spécifique existe chez ce maillon biologique. La distinction, ou non, des espèces au sein de complexes d'espèces cryptiques a pourtant une influence potentielle sur les mesures de diversité  $\beta$  par exemple, et les valeurs observées dans le chapitre 3 étaient peut être plus faibles que ce qu'elles auraient du être à cause de la non distinction des espèces au sein de ces complexes.

Une des pistes pour appréhender cette adaptation intra-spécifique pourrait être de coupler des approches taxonomiques avec de la biologie moléculaire. Si la taxonomie peut être un atout pour comprendre les modes de vie des espèces (guildes écologiques par exemple), pour mesurer un certain nombre de traits morphologiques (biovolume, ratio longueur-largeur, etc), sa complexité est aussi un frein pour la détermination des espèces. Certains complexes taxonomiques, tel que *Achnantheidium minutissimum* (Kützinger), regroupent par exemple des dizaines d'espèces (Novais et al., 2015) aux préférences écologiques très diverses et dont la distinction taxonomique entre les différentes espèces est quasiment impossible et seules quelques approches basées sur la taille des individus permettent de les différencier en sous-groupes (Van de Vijver et al., 2018). Dans ce cadre, les approches de biologie moléculaire pourraient s'avérer complémentaires pour appréhender la dynamique spatio-temporelle de la biodiversité.

## **Intérêt des démarches multi-maillons**

Enfin, les communautés diatomiques peuvent être influencées par les autres maillons biologiques vivant sur les mêmes sites. Il a par exemple été montré qu'il existait de nombreuses relations allélopathiques entre macrophytes et algues microscopiques (Erhard and Gross, 2006; Vanderstukken et al., 2011) en partie responsables de l'organisation des patrons d'assemblages des communautés (Laugaste and Reunanen, 2005). Les prédateurs brouteurs des micro-algues (micro-méiofaune) sont également un maillon influençant fortement les communautés (Timms and Moss, 1984). Les communautés phytoplanctoniques ont également un fort impact sur les communautés macrophytiques puisque la quantité de phytoplancton présent dans l'eau diminue la quantité de lumière disponible pour les macrophytes (Sand-Jensen and Søndergaard, 1981), modifiant de fait les communautés diatomiques épiphytiques (Kairesalo, 1984).

Cette complexité des relations entre groupes taxonomiques explique la relative rareté de véritables approches multi-maillons, qui apparaissent pourtant essentielles pour mieux comprendre les processus agissant sur les communautés. Il serait ainsi intéressant de mener parallèlement des expérimentations en laboratoire pour étudier les relations entre les différents maillons biologiques et la communication chimique qui existe entre eux d'un côté (Cooper and Smith, 2015; Poulson-Ellestad et al., 2014; Ruocco et al., 2018; Ternon et al., 2018), et de l'autre, mener des échantillonnages simultanés des différents maillons biologiques d'intérêts, sur le terrain et à différents niveaux trophiques.

## **Effet de masse : intérêt des démarches expérimentales**

Comme nous l'avons vu au cours de cette thèse, les processus de dispersion structurant les communautés diatomiques restent, malgré les avancées présentées, difficiles à appréhender. L'effet de masse reste un modèle théorique dont l'influence sur les communautés est bien souvent conditionnée par d'autres processus structurant ces communautés. Il serait intéressant de mener des expérimentations, dans un premier temps contrôlées (en canaux expérimentaux par exemple), puis dans un second temps en milieux naturels (le long d'un continuum amont-aval), pour chercher à mieux comprendre la dynamique spatiale et temporelle de ce processus.

Comme évoqué précédemment, il serait également intéressant de revenir sur les résultats du chapitre 3, à l'aide des résultats du chapitre 4. Dans le chapitre 3, nous avons montré qu'il existait une structuration purement spatiale des communautés de diatomées, en particulier dans les rivières fortement impactées par les nutriments, mais sans être capables de discriminer l'effet de la dispersion limitée de celui de l'effet de masse. Si la combinaison de signaux spatiaux accompagnés d'une baisse de la diversité  $\beta$  nous fait pencher pour une influence majoritaire de l'effet de masse, une manière simple de la vérifier serait de

mesurer la contribution de chaque espèce à la diversité  $\beta$  globale (Species Contribution to Beta Diversity, Legendre and De Cáceres 2013).



## Références

- AFNOR, 2007. Norme NFT 90-354 - Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD).
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723.
- Amarasekare, P., Nisbet, R.M., 2001. Spatial Heterogeneity, Source-Sink Dynamics, and the Local Coexistence of Competing Species. *The American Naturalist* 158, 572–584.
- Astorga, A., Death, R., Death, F., Paavola, R., Chakraborty, M., Muotka, T., 2014. Habitat heterogeneity drives the geographical distribution of beta diversity : the case of New Zealand stream invertebrates. *Ecology and Evolution* 4, 2693–2702.
- Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R., Muotka, T., 2012. Distance decay of similarity in freshwater communities : do macro- and microorganisms follow the same rules ? *Global Ecology and Biogeography* 21, 365–375.
- Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 19, 134–143.
- Bell, G., Lechowicz, M.J., Waterway, M.J., 2006. The Comparative Evidence Relating to Functional and Neutral Interpretations of Biological Communities. *Ecology* 87, 1378–1386.
- Bennett, J., Gilbert, B., 2016. Contrasting beta diversity among regions : how do classical and multivariate approaches compare ? *Global Ecology and Biogeography* 25, 368–377.
- Bennett, J.R., Cumming, B.F., Ginn, B.K., Smol, J.P., 2010. Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography* 19, 724–732.
- Bernard, M., Boutry, S., Lissalde, S., Guibaud, G., Saüt, M., Rebillard, J.P., Mazzella, N., 2019. Combination of passive and grab sampling strategies improves the assessment of pesticide occurrence and contamination levels in a large-scale watershed. *Science of The Total Environment* 651, 684–695.

- Blair, R.B., 2001. Birds and Butterflies Along Urban Gradients in Two Ecoregions of the United States : Is Urbanization Creating a Homogeneous Fauna ?, in : Lockwood, J.L., McKinney, M.L. (Eds.), Biotic Homogenization. Springer US, pp. 33–56.
- Blanchet, G., Legendre, P., Borcard, D., 2008a. Modelling directional spatial processes in ecological data. *Ecological Modelling* 215, 325–336.
- Blanchet, G.F., Legendre, P., Borcard, D., 2008b. Forward Selection of Explanatory Variables. *Ecology* 89, 2623–2632.
- Blüthgen, N., Simons, N.K., Jung, K., Prati, D., Renner, S.C., Boch, S., Fischer, M., Hölzel, N., Klaus, V.H., Kleinebecker, T., Tschapka, M., Weisser, W.W., Gossner, M.M., 2016. Land use imperils plant and animal community stability through changes in asynchrony rather than diversity. *Nature Communications* 7, 10697.
- Bolnick, D.I., Svanbäck, R., Araújo, M.S., Persson, L., 2007. Comparative support for the niche variation hypothesis that more generalized populations also are more heterogeneous. *Proceedings of the National Academy of Sciences* 104, 10075–10079.
- Borcard, D., Legendre, P., 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* 153, 51–68.
- Bothwell, M.L., 1989. Phosphorus-Limited Growth Dynamics of Lotic Periphytic Diatom Communities : Areal Biomass and Cellular Growth Rate Responses. *Canadian Journal of Fisheries and Aquatic Sciences* 46, 1293–1301.
- Bottin, M., 2012. Structure des assemblages de diatomées benthiques en rivière : l'environnement explique-t-il tout ? : processus écologiques et développement méthodologiques. Bordeaux 1.
- Bottin, M., Soininen, J., Alard, D., Rosebery, J., 2016. Diatom Cooccurrence Shows Less Segregation than Predicted from Niche Modeling. *PLOS ONE* 11, e0154581.
- Bottin, M., Soininen, J., Ferrol, M., Tison-Rosebery, J., 2014. Do spatial patterns of benthic diatom assemblages vary across regions and years ? *Freshwater Science* 33, 402–416.
- Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26, 211–252.
- Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27, 437–448.

- Brown, B.L., Swan, C.M., 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology* 79, 571–580.
- Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O., Patrick, C., 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society* 30, 310–327.
- Chase, J.M., 2010. Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. *Science* , 1187820.
- Chase, J.M., 2014. Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science* 25, 319–322.
- Chase, J.M., Amarasekare, P., Cottenie, K., Gonzalez, A., Holt, R.D., Holyoak, M., Hoopes, M.F., Leibold, M.A., Loreau, M., Mouquet, N., Shurin, J.B., Tilman, D., 2005. Competing theories for competitive metacommunities. *Metacommunities : spatial dynamics and ecological communities* , 335–354.
- Chase, J.M., Kraft, N.J.B., Smith, K.G., Vellend, M., Inouye, B.D., 2011. Using null models to disentangle variation in community dissimilarity from variation in alpha-diversity. *Ecosphere* 2, 1–11.
- Chase, J.M., Leibold, M.A., 2003. *Ecological Niches : Linking Classical and Contemporary Approaches*. University of Chicago Press.
- Chase, J.M., Myers, J.A., 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B : Biological Sciences* 366, 2351–2363.
- Chételat, J., Pick, F.R., Morin, A., Hamilton, P.B., 1999. Periphyton biomass and community composition in rivers of different nutrient status. *Canadian Journal of Fisheries and Aquatic Sciences* 56, 560–569.
- Clark, J.S., Bell, D., Chu, C., Courbaud, B., Dietze, M., Hersh, M., HilleRisLambers, J., Ibáñez, I., LaDeau, S., McMahon, S., Metcalf, J., Mohan, J., Moran, E., Pangle, L., Pearson, S., Salk, C., Shen, Z., Valle, D., Wyckoff, P., 2010. High-dimensional coexistence based on individual variation : a synthesis of evidence. *Ecological Monographs* 80, 569–608.
- Clark, J.S., Dietze, M., Chakraborty, S., Agarwal, P.K., Ibanez, I., LaDeau, S., Wolosin, M., 2007. Resolving the biodiversity paradox. *Ecology Letters* 10, 647–659.
- Clavel, J., Julliard, R., Devictor, V., 2011. Worldwide decline of specialist species : toward

- a global functional homogenization ? *Frontiers in Ecology and the Environment* 9, 222–228.
- Cliff, N., 1993. Dominance statistics : Ordinal analyses to answer ordinal questions. *Psychological Bulletin* , 494–509.
- Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H.C., Losos, E., Hubbell, S.P., 2002. Beta-Diversity in Tropical Forest Trees. *Science* 295, 666–669.
- Conroy, M.J., Noon, B.R., 1996. Mapping of Species Richness for Conservation of Biological Diversity : Conceptual and Methodological Issues. *Ecological Applications* 6, 763–773.
- Cooper, M.B., Smith, A.G., 2015. Exploring mutualistic interactions between microalgae and bacteria in the omics age. *Current Opinion in Plant Biology* 26, 147–153.
- Coste, M., 1982. Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse .
- Coste, M., Boutry, S., Tison-Rosebery, J., Delmas, F., 2009. Improvements of the Biological Diatom Index (BDI) : Description and efficiency of the new version (BDI-2006). *Ecological Indicators* 9, 621–650.
- Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* 8, 1175–1182.
- Cottenie, K., Michels, E., Nuytten, N., De Meester, L., 2003. Zooplankton Metacommunity Structure : Regional Vs. Local Processes in Highly Interconnected Ponds. *Ecology* 84, 991–1000.
- De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W., Declerck, S.A.J., 2012. Body size and dispersal mode as key traits determining meta-community structure of aquatic organisms. *Ecology Letters* 15, 740–747.
- DeFries, R.S., Foley, J.A., Asner, G.P., 2004. Land-use choices : balancing human needs and ecosystem function. *Frontiers in Ecology and the Environment* 2, 249–257.
- Dolédec, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche Separation in Community Analysis : A New Method. *Ecology* 81, 2914–2927.
- Dong, X., Li, B., He, F., Gu, Y., Sun, M., Zhang, H., Tan, L., Xiao, W., Liu, S., Cai, Q., 2016.

- Flow directionality, mountain barriers and functional traits determine diatom metacom-  
munity structuring of high mountain streams. *Scientific Reports* 6.
- Dray, S., Blanchet, G., Borcard, D., Clappe, S., Jombart, G.G.T., Larocque, G., Legendre,  
P., Madi, N., Wagner, H.H., 2018. *adespatial* : Moran's eigenvector maps and related  
methods for the spatial multiscale analysis of ecological communities. R package ver-  
sion 0.1-1.
- Ehrlén, J., Eriksson, O., 2000. Dispersal Limitation and Patch Occupancy in Forest Herbs.  
*Ecology* 81, 1667–1674.
- E Kroos, J., Heliölä, J., Kuussaari, M., 2010. Homogenization of lepidopteran communities  
in intensively cultivated agricultural landscapes. *Journal of Applied Ecology* 47, 459–  
467.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees.  
*Journal of Animal Ecology* 77, 802–813.
- Erhard, D., Gross, E.M., 2006. Allelopathic activity of *Elodea canadensis* and *Elodea*  
*nuttallii* against epiphytes and phytoplankton. *Aquatic Botany* 85, 203–211.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller,  
M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner,  
M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission.  
*Reviews of Geophysics* 45.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors  
in conservation presence/absence models. *Environmental Conservation* 24, 38–49.
- Finlay, B.J., 2002. Global Dispersal of Free-Living Microbial Eukaryote Species. *Science*  
296, 1061–1063.
- Fridley, J.D., Vandermaast, D.B., Kuppinger, D.M., Manthey, M., Peet, R.K., 2007. Co-  
occurrence based assessment of habitat generalists and specialists : a new approach  
for the measurement of niche width. *Journal of Ecology* 95, 707–722.
- Gabriel, D., Roschewitz, I., Tschardtke, T., Thies, C., 2006. Beta Diversity at Different  
Spatial Scales : Plant Communities in Organic and Conventional Agriculture. *Ecological*  
*Applications* 16, 2011–2021.
- Gilbert, B., Bennett, J.R., 2010. Partitioning variation in ecological communities : do the  
numbers add up? *Journal of Applied Ecology* 47, 1071–1082.
- Goldenberg Vilar, A., van Dam, H., van Loon, E.E., Vonk, J.A., van Der Geest, H.G.,

- Admiraal, W., 2014. Eutrophication decreases distance decay of similarity in diatom communities. *Freshwater Biology* 59, 1522–1531.
- Gossner, M.M., Lewinsohn, T.M., Kahl, T., Grassein, F., Boch, S., Prati, D., Birkhofer, K., Renner, S.C., Sikorski, J., Wubet, T., Arndt, H., Baumgartner, V., Blaser, S., Blüthgen, N., Börschig, C., Buscot, F., Diekötter, T., Jorge, L.R., Jung, K., Keyel, A.C., Klein, A.M., Klemmer, S., Krauss, J., Lange, M., Müller, J., Overmann, J., Pašalić, E., Penone, C., Perović, D.J., Purschke, O., Schall, P., Socher, S.A., Sonnemann, I., Tschapka, M., Tscharnkte, T., Türke, M., Venter, P.C., Weiner, C.N., Werner, M., Wolters, V., Wurst, S., Westphal, C., Fischer, M., Weisser, W.W., Allan, E., 2016. Land-use intensification causes multitrophic homogenization of grassland communities. *Nature* 540, 266–269.
- Grant, E.H.C., Lowe, W.H., Fagan, W.F., 2007. Living in the branches : population dynamics and ecological processes in dendritic networks. *Ecology Letters* 10, 165–175.
- Griffith, D.M., Veech, J.A., Marsh, C.J., 2016. Cooccur : probabilistic species co-occurrence analysis in R. *J Stat Softw* 69, 1–17.
- Grinnell, J., 1917. The niche-relationships of the california thrasher. *The Auk* 34, 427–433.
- Göthe, E., Angeler, D.G., Gottschalk, S., Löfgren, S., Sandin, L., 2013. The Influence of Environmental, Biotic and Spatial Factors on Diatom Metacommunity Structure in Swedish Headwater Streams. *PLOS ONE* 8, e72237.
- Gutiérrez-Cánovas, C., Millán, A., Velasco, J., Vaughan, I.P., Ormerod, S.J., 2013. Contrasting effects of natural and anthropogenic stressors on beta diversity in river organisms. *Global Ecology and Biogeography* 22, 796–805.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136 : A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100–108.
- Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability ? *Ecography* 35, 276–288.
- Heino, J., Grönroos, M., 2017. Exploring species and site contributions to beta diversity in stream insect assemblages. *Oecologia* 183, 151–160.
- Heino, J., Grönroos, M., Ilmonen, J., Karhu, T., Niva, M., Paasivirta, L., 2013. Environmental heterogeneity and beta diversity of stream macroinvertebrate communities at intermediate spatial scales. *Freshwater Science* 32, 142–154.
- 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. *Freshwater Biology* 60, 845–869.

- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology* 25, 1965–1978.
- Hoagland, K.D., Rosowski, J.R., Gretz, M.R., Roemer, S.C., 1993. Diatom Extracellular Polymeric Substances : Function, Fine Structure, Chemistry, and Physiology. *Journal of Phycology* 29, 537–566.
- Holyoak, M., Leibold, M.A., Holt, R.D., 2005. *Metacommunities : Spatial Dynamics and Ecological Communities*. University of Chicago Press.
- Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- Huston, M.A., 1994. *Biological Diversity : The Coexistence of Species*. Cambridge University Press.
- Hutchinson, G., 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22, 145–159.
- Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis : A Comparison of Heuristical and Statistical Approaches. *Ecology* 74, 2204–2214.
- Jamoneau, A., Passy, S.I., Soininen, J., Lebourcier, T., Tison-Rosebery, J., 2018. Beta diversity of diatom species and ecological guilds : Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology* 63, 62–73.
- Jenkins, D.G., Brescacin, C.R., Duxbury, C.V., Elliott, J.A., Evans, J.A., Grablow, K.R., Hillegass, M., Lyon, B.N., Metzger, G.A., Olandese, M.L., Pepe, D., Silvers, G.A., Suresch, H.N., Thompson, T.N., Trexler, C.M., Williams, G.E., Williams, N.C., Williams, S.E., 2007. Does size matter for dispersal distance ? *Global Ecology and Biogeography* 16, 415–425.
- Johnson, R.K., Furse, M.T., Hering, D., Sandin, L., 2007. Ecological relationships between stream communities and spatial scale : implications for designing catchment-level monitoring programmes. *Freshwater Biology* 52, 939–958.
- JORF, 2015. Arrêté du 27 juillet 2015 modifiant l'arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement. *Journal Officiel de la République Française* .
- Jung, V., Albert, C.H., Violle, C., Kunstler, G., Loucougaray, G., Spiegelberger, T., 2014.

- Intraspecific trait variability mediates the response of subalpine grassland communities to extreme drought events. *Journal of Ecology* 102, 45–53.
- Jung, V., Violle, C., Mondy, C., Hoffmann, L., Muller, S., 2010. Intraspecific variability and trait-based community assembly. *Journal of Ecology* 98, 1134–1140.
- Kairesalo, T., 1984. The Seasonal Succession of Epiphytic Communities within an *Equisetum fluvatile* L. Stand in Lake Pääjärvi, Southern Finland. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 69, 475–505.
- Karp, D.S., Rominger, A.J., Zook, J., Ranganathan, J., Ehrlich, P.R., Daily, G.C., 2012. Intensive agriculture erodes beta-diversity at large scales. *Ecology Letters* 15, 963–970.
- Kaspari, M., Stevenson, B.S., Shik, J., Kerekes, J.F., 2010. Scaling community structure : how bacteria, fungi, and ant taxocenes differentiate along a tropical forest floor. *Ecology* 91, 2221–2226.
- Kodinariya, T.M., Makwana, P.R., 2013. Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies* 1, 90–95.
- Kraft, N.J.B., Comita, L.S., Chase, J.M., Sanders, N.J., Swenson, N.G., Crist, T.O., Stegen, J.C., Vellend, M., Boyle, B., Anderson, M.J., Cornell, H.V., Davies, K.F., Freestone, A.L., Inouye, B.D., Harrison, S.P., Myers, J.A., 2011. Disentangling the Drivers of Beta-Diversity Along Latitudinal and Elevational Gradients. *Science* 333, 1755–1758.
- Lack, D., 1976. *Island biology ; illustrated by the land birds of Jamaica* .
- Laliberté, E., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3232–3237.
- Landeiro, V.L., Magnusson, W.E., Melo, A.S., Espírito-Santo, H.M.V., Bini, L.M., 2011. Spatial eigenfunction analyses in stream networks : do watercourse and overland distances produce different results ? *Freshwater Biology* 56, 1184–1192.
- Larsen, S., Ormerod, S.J., 2014. Anthropogenic modification disrupts species co-occurrence in stream invertebrates. *Global Change Biology* 20, 51–60.
- Laugaste, R., Reunanen, M., 2005. The composition and density of epiphyton on some macrophyte species in the partly meromictic Lake Verevi, in : Ott, I., Kõiv, T. (Eds.), *Lake Verevi, Estonia — A Highly Stratified Hypertrophic Lake*. Springer Netherlands, Dordrecht. *Developments in Hydrobiology*, pp. 137–150.

- Lecointe, C., Coste, M., Prygiel, J., 1993. Omnidia : software for taxonomy, calculation of diatom indices and inventories management. *Hydrobiologia* 269-270, 509–513.
- Legendre, P., Borcard, D., Peres-Neto, P.R., 2005. Analyzing Beta Diversity : Partitioning the Spatial Variation of Community Composition Data. *Ecological Monographs* 75, 435–450.
- Legendre, P., Borcard, D., Peres-Neto, P.R., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3238–3244.
- Legendre, P., De Cáceres, M., 2013. Beta diversity as the variance of community data : dissimilarity coefficients and partitioning. *Ecology Letters* 16, 951–963.
- Legendre, P., Legendre, L.F.J., 2012. *Numerical Ecology*. Elsevier.
- 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., Gonzalez, A., 2004. The metacommunity concept : a framework for multi-scale community ecology. *Ecology Letters* 7, 601–613.
- Lenoir, A., Coste, M., 1996. Development of a practical diatom index of overall water quality applicable to the French national water board network. Use of algae for monitoring rivers , 29–43.
- Letten, A.D., Ke, P.J., Fukami, T., 2017. Linking modern coexistence theory and contemporary niche theory. *Ecological Monographs* 87, 161–177.
- Liu, J., Soininen, J., Han, B.P., Declerck, S.A.J., 2013. Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of Biogeography* 40, 2238–2248.
- Mac Nally, R., Fleishman, E., Bulluck, L.P., Betrus, C.J., 2004. Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography* 31, 917–929.
- MacArthur, R., Levins, R., 1967. The limiting similarity, convergence, and divergence of coexisting species. *The american naturalist* 101, 377–385.
- MacArthur, R.H., Wilson, E.O., 1967. *The theory of island biogeography*. volume 1. Princeton university press.
- Mann, D.G., Vanormelingen, P., 2013. An Inordinate Fondness ? The Number, Distributions, and Origins of Diatom Species. *Journal of Eukaryotic Microbiology* 60, 414–420.

- Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., Horner-Devine, M.C., 2011. Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences* 108, 7850–7854.
- McGill, B.J., Dornelas, M., Gotelli, N.J., Magurran, A.E., 2015. Fifteen forms of biodiversity trend in the Anthropocene. *Trends in Ecology & Evolution* 30, 104–113.
- McKinney, M.L., 2006. Urbanization as a major cause of biotic homogenization. *Biological Conservation* 127, 247–260.
- McKinney, M.L., Lockwood, J.L., 1999. Biotic homogenization : a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* 14, 450–453.
- Mouquet, N., Loreau, M., 2003. Community Patterns in Source Sink Metacommunities. *The American Naturalist* 162, 544–557.
- Myers, J.A., Chase, J.M., Jiménez, I., Jørgensen, P.M., Araujo-Murakami, A., Paniagua-Zambrana, N., Seidel, R., 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters* 16, 151–157.
- Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M.J., Feldman, A., Garon, M., Harrison, M.L.K., Alhusseini, T., Ingram, D.J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Correia, D.L.P., Martin, C.D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H.R.P., Purves, D.W., Robinson, A., Simpson, J., Tuck, S.L., Weiher, E., White, H.J., Ewers, R.M., Mace, G.M., Scharlemann, J.P.W., Purvis, A., 2015. Global effects of land use on local terrestrial biodiversity. *Nature* 520, 45–50.
- Ng, I.S.Y., Carr, C.M., Cottenie, K., 2009. Hierarchical zooplankton metacommunities : distinguishing between high and limiting dispersal mechanisms. *Hydrobiologia* 619, 133–143.
- Noss, R.F., 1983. A Regional Landscape Approach to Maintain Diversity. *BioScience* 33, 700–706.
- Novais, M.H., Jüttner, I., Van de Vijver, B., Morais, M.M., Hoffmann, L., Ector, L., 2015. Morphological variability within the *Achnanthydium minutissimum* species complex (Bacillariophyta) : comparison between the type material of *Achnanthes minutissima* and related taxa, and new freshwater *Achnanthydium* species from Portugal. *Phytotaxa* 224, 101–139.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson,

- G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. *vegan* : Community ecology package. R package version , 117–118.
- Olden, J., Poff, N., 2003. Toward a Mechanistic Understanding and Prediction of Biotic Homogenization. *The American Naturalist* 162, 442–460.
- Olden, J.D., Poff, N.L., 2004. Ecological Processes Driving Biotic Homogenization : Testing a Mechanistic Model Using Fish Faunas. *Ecology* 85, 1867–1875.
- Olden, J.D., Rooney, T.P., 2006. On defining and quantifying biotic homogenization. *Global Ecology and Biogeography* 15, 113–120.
- Ozkan, K., Svenning, J.C., Jeppesen, E., Webb, T., 2013. Environmental species sorting dominates forest bird community assembly across scales. *Journal of Animal Ecology* 82, 266–274.
- Pajunen, V., Luoto, M., Soininen, J., 2016. Climate is an important driver for stream diatom distributions. *Global Ecology and Biogeography* 25, 198–206.
- Palmer, M.A., Allan, J.D., Butman, C.A., 1996. Dispersal as a regional process affecting the local dynamics of marine and stream benthic invertebrates. *Trends in ecology & evolution* 11, 322–326.
- Passy, S.I., 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquatic Botany* 86, 171–178.
- Passy, S.I., 2009. The relationship between local and regional diatom richness is mediated by the local and regional environment. *Global Ecology and Biogeography* 18, 383–391.
- Pella, H., Lejot, J., Lamouroux, N., Snelder, T., 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. *Géomorphologie : relief, processus, environnement* 18, 317–336.
- Pélissier, R., Couteron, P., Dray, S., 2008. Analyzing or Explaining Beta Diversity ? Comment. *Ecology* 89, 3227–3232.
- Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., Schmutz, S., 2006. Assessing river biotic condition at a continental scale : a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology* 43, 70–80.
- Potapova, M.G., Charles, D.F., 2002. Benthic diatoms in USA rivers : distributions along spatial and environmental gradients. *Journal of Biogeography* 29, 167–187.

- Poulson-Ellestad, K.L., Jones, C.M., Roy, J., Viant, M.R., Fernández, F.M., Kubanek, J., Nunn, B.L., 2014. Metabolomics and proteomics reveal impacts of chemically mediated competition on marine plankton. *Proceedings of the National Academy of Sciences* , 201402130.
- Pound, K., Lawrence, G.B., Passy, S.I., 2018. Beta diversity response to stress severity and heterogeneity in sensitive versus tolerant stream diatoms. *Diversity and Distributions* in press.
- Prygiel, J., Coste, M., 2000. Guide méthodologique pour la mise en oeuvre de l'indice biologique diatomées n° 90-354. Agences de l'Eau, Ministère de l'Aménagement du Territoire et de l'Environnement, Cemagref .
- Prygiel, J., Coste, M., Bukowska, J., 1999. Review of the major diatom-based techniques for the quality assessment of rivers-State of the art in Europe. *Use of algae for monitoring rivers* 3, 224–238.
- R Core Team, 2017. R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.
- R Core Team, 2018. R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Romano, J., Kromrey, J.D., Coraggio, J., Skowronek, J., 2016. Appropriate statistics for ordinal level data : Should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys. *Annual Meeting of the Florida Association of Institutional Research* , 1–33.
- Ruocco, N., Costantini, S., Zupo, V., Lauritano, C., Caramiello, D., Ianora, A., Budillon, A., Romano, G., Nuzzo, G., D'Ippolito, G., Fontana, A., Costantini, M., 2018. Toxigenic effects of two benthic diatoms upon grazing activity of the sea urchin : morphological, metabolomic and de novo transcriptomic analysis. *Scientific Reports* 8, 5622.
- Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M., Wall, D.H., 2000. Global Biodiversity Scenarios for the Year 2100. *Science* 287, 1770–1774.
- Sand-Jensen, K., Søndergaard, M., 1981. Phytoplankton and Epiphyte Development and Their Shading Effect on Submerged Macrophytes in Lakes of Different Nutrient Status. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 66, 529–552.
- Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Sam-

- son, F.B., 2002. Predicting species occurrences : issues of scale and accuracy. Island press ed.
- Shmida, A., Wilson, M.V., 1985. Biological Determinants of Species Diversity. *Journal of Biogeography* 12, 1–20.
- Siepielski, A.M., McPeck, M.A., 2010. On the evidence for species coexistence : a critique of the coexistence program. *Ecology* 91, 3153–3164.
- Smucker, N.J., Vis, M.L., 2011a. Diatom biomonitoring of streams : Reliability of reference sites and the response of metrics to environmental variations across temporal scales. *Ecological Indicators* 11, 1647–1657.
- Smucker, N.J., Vis, M.L., 2011b. Spatial factors contribute to benthic diatom structure in streams across spatial scales : Considerations for biomonitoring. *Ecological Indicators* 11, 1191–1203.
- Socolar, J.B., Gilroy, J.J., Kunin, W.E., Edwards, D.P., 2016. How Should Beta-Diversity Inform Biodiversity Conservation ? *Trends in Ecology & Evolution* 31, 67–80.
- Soininen, J., 2007. Environmental and Spatial Control of Freshwater Diatoms - a Review. *Diatom Research* 22, 473–490.
- Soininen, J., Heino, J., Wang, J., 2018. A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. *Global Ecology and Biogeography* 27, 96–109.
- Soininen, J., Jamoneau, A., Rosebery, J., Passy, S.I., 2016. Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography* 25, 940–950.
- Soininen, J., Lennon, J.J., Hillebrand, H., 2007. A Multivariate Analysis of Beta Diversity across Organisms and Environments. *Ecology* 88, 2830–2838.
- Solar, R.R.C., Barlow, J., Ferreira, J., Berenguer, E., Lees, A.C., Thomson, J.R., Louzada, J., Maués, M., Moura, N.G., Oliveira, V.H.F., Chaul, J.C.M., Schoereder, J.H., Vieira, I.C.G., Mac Nally, R., Gardner, T.A., 2015. How pervasive is biotic homogenization in human-modified tropical forest landscapes ? *Ecology Letters* 18, 1108–1118.
- Sotherton, N.W., 1998. Land use changes and the decline of farmland wildlife : An appraisal of the set-aside approach. *Biological Conservation* 83, 259–268.
- Stevenson, R.J., Rier, S.T., Riseng, C.M., Schultz, R.E., Wiley, M.J., 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance

- regimes and with applications for developing nutrient criteria, in : *Advances in Algal Biology : A Commemoration of the Work of Rex Lowe*. Springer, Dordrecht. *Developments in Hydrobiology*, pp. 149–165.
- Svenning, J.C., Normand, S., Skov, F., 2008. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography* 31, 316–326.
- Tecchio, S., Chaalali, A., Raoux, A., Tous Rius, A., Lequesne, J., Girardin, V., Lassalle, G., Cachera, M., Riou, P., Lobry, J., Dauvin, J.C., Niquil, N., 2016. Evaluating ecosystem-level anthropogenic impacts in a stressed transitional environment : The case of the Seine estuary. *Ecological Indicators* 61, 833–845.
- Tello, J.S., Myers, J.A., Macía, M.J., Fuentes, A.F., Cayola, L., Arellano, G., Loza, M.I., Torrez, V., Cornejo, M., Miranda, T.B., Jørgensen, P.M., 2015. Elevational gradients in beta-diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLOS ONE* 10, e0121458.
- Ternon, E., Pavaux, A.S., Marro, S., Thomas, O.P., Lemée, R., 2018. Allelopathic interactions between the benthic toxic dinoflagellate *Ostreopsis* cf. *ovata* and a co-occurring diatom. *Harmful Algae* 75, 35–44.
- Timms, R.M., Moss, B., 1984. Prevention of growth of potentially dense phytoplankton populations by zooplankton grazing, in the presence of zooplanktivorous fish, in a shallow wetland ecosystem. *Limnology and Oceanography* 29, 472–486.
- Tison, J., Park, Y.S., Coste, M., Wasson, J.G., Ector, L., Rimet, F., Delmas, F., 2005. Typology of diatom communities and the influence of hydro-ecoregions : A study on the French hydrosystem scale. *Water Research* 39, 3177–3188.
- Tonkin, J.D., Altermatt, F., Finn, D.S., Heino, J., Olden, J.D., Pauls, S.U., Lytle, D.A., 2018. The role of dispersal in river network metacommunities : Patterns, processes, and pathways. *Freshwater Biology* 63, 141–163.
- Tonkin, J.D., Sundermann, A., Jähnig, S.C., Haase, P., 2015. Environmental Controls on River Assemblages at the Regional Scale : An Application of the Elements of Metacommunity Structure Framework. *PLOS ONE* 10, e0135450.
- Tudesque, L., Tisseuil, C., Lek, S., 2014. Scale-dependent effects of land cover on water physico-chemistry and diatom-based metrics in a major river system, the Adour-Garonne basin (South Western France). *Science of The Total Environment* 466-467, 47–55.
- Tuomisto, H., 2010a. A diversity of beta diversities : straightening up a concept gone awry.

- Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33, 2–22.
- Tuomisto, H., 2010b. A diversity of beta diversities : straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33, 23–45.
- Tuomisto, H., Ruokolainen, K., 2006. Analyzing or Explaining Beta Diversity ? Understanding the Targets of Different Methods of Analysis. *Ecology* 87, 2697–2708.
- Tuomisto, H., Ruokolainen, K., Yli-Halla, M., 2003. Dispersal, Environment, and Floristic Variation of Western Amazonian Forests. *Science* 299, 241–244.
- Ulrich, W., Baselga, A., Kusumoto, B., Shiono, T., Tuomisto, H., Kubota, Y., Peres-Neto, P., 2017. The tangled link between beta- and gamma-diversity : a Narcissus effect weakens statistical inferences in null model analyses of diversity patterns. *Global Ecology and Biogeography* 26, 1–5.
- Underwood, A.J., Chapman, M.G., 1996. Scales of spatial patterns of distribution of intertidal invertebrates. *Oecologia* 107, 212–224.
- Union Européenne, 2006. Corine Land Cover. SOeS .
- Vallejos, M.A.V., Padial, A.A., Vitule, J.R.S., 2016. Human-Induced Landscape Changes Homogenize Atlantic Forest Bird Assemblages through Nested Species Loss. *PLOS ONE* 11, e0147058.
- Van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Netherland Journal of Aquatic Ecology* 28, 117–133.
- Vanderstukken, M., Mazzeo, N., Van Colen, W., Declerck, S.J., Muylaert, K., 2011. Biological control of phytoplankton by the subtropical submerged macrophytes *Egeria densa* and *Potamogeton illinoensis* : a mesocosm study. *Freshwater Biology* 56, 1837–1849.
- Vanormelingen, P., Verleyen, E., Vyverman, W., 2008. The diversity and distribution of diatoms : from cosmopolitanism to narrow endemism. *Biodiversity and Conservation* 17, 393–405.
- Verleyen, E., Vyverman, W., Sterken, M., Hodgson, D.A., De Wever, A., Juggins, S., Van de Vijver, B., Jones, V.J., Vanormelingen, P., Roberts, D., Flower, R., Kilroy, C., Souffreau, C., Sabbe, K., 2009. The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos* 118, 1239–1249.
- Viana, D.S., Figuerola, J., Schwenk, K., Manca, M., Hobæk, A., Mjelde, M., Preston, C.D.,

- Gornall, R.J., Croft, J.M., King, R.A., Green, A.J., Santamaría, L., 2016. Assembly mechanisms determining high species turnover in aquatic communities over regional and continental scales. *Ecography* 39, 281–288.
- Van de Vijver, B., Mertens, A., Wetzel, C., Ector, L., 2018. *Achnanthidium minutissimum*. 5ème workshop taxonomique de l'ADLAF .
- Violle, C., Enquist, B.J., McGill, B.J., Jiang, L., Albert, C.H., Hulshof, C., Jung, V., Messier, J., 2012. The return of the variance : intraspecific variability in community ecology. *Trends in Ecology & Evolution* 27, 244–252.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997. Human Domination of Earth's Ecosystems. *Science* 277, 494–499.
- Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. *Nature* 467, 555.
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Vijver, B.V., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P., Wever, A.D., 2007. Historical Processes Constrain Patterns in Global Diatom Diversity. *Ecology* 88, 1924–1931.
- Whitlock, R., Grime, J.P., Booth, R., Burke, T., 2007. The role of genotypic diversity in determining grassland community structure under constant environmental conditions. *Journal of Ecology* 95, 895–907.
- Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30, 279–338.
- Whittaker, R.H., 1962. Classification of natural communities. *The Botanical Review* 28, 1–239.
- Whittaker, R.H., 1972. Evolution and Measurement of Species Diversity. *Taxon* 21, 213–251.
- Willis, K.J., Whittaker, R.J., 2002. Species Diversity–Scale Matters. *Science* 295, 1245–1248.
- Wilson, M.V., Shmida, A., 1984. Measuring Beta Diversity with Presence-Absence Data. *Journal of Ecology* 72, 1055–1064.
- Zelený, D., 2009. Cooccurrence based assessment of species habitat specialization is affected by the size of species pool : reply to. *Journal of Ecology* 97, 10–17.

Zelinka, M., Marvan, P., 1961. Zur prazisierung der biologischen klassifikation der reinheid fließender gewasser. Arch. Hydrobiol. 57, 389–407.