



HAL
open science

Mise à jour de l'Indice Biologique Diatomées : un enjeu pour la gestion des cours d'eau français

Léonard Henry, Sébastien Boutry, Juliette Tison-Rosebery

► To cite this version:

Léonard Henry, Sébastien Boutry, Juliette Tison-Rosebery. Mise à jour de l'Indice Biologique Diatomées : un enjeu pour la gestion des cours d'eau français. INRAE - EABX. 2023. hal-04645786

HAL Id: hal-04645786

<https://hal.inrae.fr/hal-04645786v1>

Submitted on 12 Jul 2024

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.

Mise à jour de l'Indice Biologique Diatomées : un enjeu pour la gestion des cours d'eau français

Léonard HEINRY, Sébastien BOUTRY, Juliette ROSEBERY

Août 2023

Document final

En partenariat avec les experts diatomées de
l'Office Français de la Biodiversité :

Valérie PEETERS, David HEUDRE, Rémy CHAUAUX,
Jean-Marc VOUTERS, Servanne QUINIOU

Avec le soutien de

Contexte de programmation et de réalisation

Ce rapport a été réalisé dans le cadre du programme scientifique et technique AQUAREF pour l'année 2023, au titre de l'action « Mise à jour des profils indiciaires de l'Indice Biologique Diatomées », thème « B ».

Auteur (s) :

Léonard Henry
Inrae
leonard.heinry@gmail.com

Juliette Rosebery
Inrae
juliette.Rosebery@inrae.fr

Sébastien Boutry
Inrae
sebastien.Boutry@inrae.fr

Vérification du document :

Rémi Buchet
Ifremer centre de Nantes
remi.buchet@ifremer.fr

Claire Rollet
Ifremer Station de Dinard
Claire.rollet@ifremer.fr

Les correspondants

OFB : Nicolas Hette-Tronquart ; nicolas.hette-tronquart@ofb.gouv.fr

Référence du document : Léonard Henry, Sébastien Boutry, Juliette Rosebery (2023). Mise à jour de l'Indice Biologique Diatomées : un enjeu pour la gestion des cours d'eau français - Rapport AQUAREF 2023 - 22p. + annexes

Droits d'usage :	<i>Accès libre</i>
Couverture géographique :	<i>Nationale</i>
Niveau géographique :	<i>National</i>
Niveau de lecture :	<i>Professionnels, experts</i>
Nature de la ressource :	<i>Document</i>

1	INTRODUCTION.....	1
2	MÉTHODE.....	2
2.1	Origine des données exploitées.....	2
2.2	Taxons cibles	2
2.2.1	Liste proposée par les experts diatomistes	2
2.2.2	Transcodage	3
2.2.3	Base de données de travail	4
2.3	Développement d'une méthode pour associer des profils écologiques aux taxons cibles	5
2.3.1	Construction d'une base de données de comparaison	5
2.3.2	Taxons indicateurs selon la méthode IndVal	6
2.3.3	Développement de l'approche Indval pour identifier les taxons aux distributions équivalentes	7
2.3.4	Méthodologie générale d'attribution d'un profil écologique	8
2.4	Impact de l'intégration des taxons cibles sur l'évaluation de l'état écologique	9
2.5	Développement d'une application de visualisation de la base NAIADES	9
3	RÉSULTATS / DISCUSSION	12
3.1	Base de données de travail	12
3.2	Attribution d'un profil écologique aux taxons cibles.....	15
3.3	Impact des nouveaux taxons contributifs sur les notes d'IBD	17
4	CONCLUSION.....	20
5	RÉFÉRENCES	21
6	ANNEXES	23

Liste des annexes

Annexe 1 : Tableau des occurrences des taxons cibles dans les relevés floristiques

Mise à jour de l'Indice Biologique Diatomées : un enjeu pour la gestion des cours d'eau français

Léonard HEINRY, Sébastien BOUTRY, Juliette ROSEBERY

RÉSUMÉ

Depuis plus de 30 ans, les diatomées benthiques sont utilisées comme bio-indicateurs pour évaluer l'état écologique des cours d'eau grâce à l'Indice Biologique Diatomées, aussi appelé IBD. Néanmoins, les récentes avancées de la taxonomie de ces microalgues et l'apparition d'espèces invasives modifient la composition des communautés de diatomées, rendant en partie obsolète la liste des espèces participant au calcul de l'IBD. Par une méthode d'analyse d'espèces indicatrices, nous avons révisé/déterminé le profil écologique de taxons considérés par les experts diatomistes de l'Office Français de la Biodiversité comme susceptibles d'intervenir dorénavant dans le calcul de l'IBD. Grâce à leur collaboration, nous avons pu caractériser le profil de 56 taxons qui entreront dans le calcul de l'indice d'ici le début de l'année 2024. Face à la vitesse avec laquelle la taxonomie évolue, ce rapport souligne la nécessité de réviser régulièrement la liste des taxons indicateurs de l'IBD ainsi que leurs profils écologiques, afin de permettre une évaluation et une gestion optimale de la qualité écologique des milieux aquatiques.

Mots clés :

Diatomées, IBD, taxon indicateur, profil écologique

Biological Diatom Index revision: A challenge for French rivers monitoring

Léonard HEINRY, Sébastien BOUTRY, Juliette ROSEBERY

ABSTRACT

Over the past three decades, the Biological Diatom Index (BDI) has contributed to the assessment of freshwater ecological status using benthic diatoms as bio-indicators. However, recent advances in the taxonomy of these micro-algae coupled with the growing influx of invasive species has led to shifts in community composition. Thus, the list of indicator taxa contributing to the calculation of the BDI is likely to be no longer valid. Using a species indicator analysis, we revised or determined the ecological profile of taxa that could now participate to the BDI computation. Through a collaborative effort involving diatom experts from the French biodiversity office, we successfully assigned an ecological profile to 56 taxa that will now participate to the BDI calculation from 2024. Considering the frequent taxonomy changes, our work supports the necessity for periodic reviews of the BDI indicator taxa list, so as to allow an assessment and optimal management of the ecological quality of aquatic environments

Key words:

Diatoms, BDI, indicator taxa, ecological profiles

1 INTRODUCTION

La gestion des milieux aquatiques d'eau douce représente un enjeu majeur pour la plupart des pays européens. Des cours d'eau de bonne qualité facilitent l'accès à l'eau potable et favorisent les activités récréatives qui stimulent le tourisme et contribuent à maintenir l'économie locale. Les services écosystémiques fournis par les fleuves, les rivières, les ruisseaux, les lacs et les étangs en font des ressources naturelles précieuses, au cœur de l'économie, de la culture et de l'éducation [1]. Paradoxalement, ces milieux comptent parmi les écosystèmes les plus dégradés. L'élargissement des zones urbaines et les nombreuses activités humaines modifient l'état des berges, fragmentent les habitats et rejettent de nombreuses substances chimiques susceptibles d'impacter de manière durable les écosystèmes aquatiques [2]. L'enrichissement des milieux en nutriments et l'accumulation de produits toxiques ont un effet direct sur la disponibilité des ressources et la biodiversité ([3], [4], [5]). En plus d'altérer la stabilité des écosystèmes, ces bouleversements créent des conditions favorables à la colonisation de nouvelles espèces invasives qui accélèrent le déclin de certaines espèces endémiques [6]. Face à ces multiples pressions, l'état écologique des cours d'eau est devenu une inquiétude d'ordre prioritaire en Europe.

La Directive Cadre sur l'Eau¹ (DCE) a placé les organismes vivants au cœur de cette évaluation de l'état écologique, notamment les diatomées, microalgues benthiques utilisées comme indicateurs biologiques depuis plus de 30 ans. La caractérisation de la réponse des différentes espèces à des conditions environnementales précises permet de garantir un suivi régulier de l'acidification, de l'état trophique, des variations de température et des variations de conductivité des milieux d'eau douce [7] [8]. L'Indice Biologique Diatomées (IBD) initialement développé en 1996, normalisé en 2000 et révisé en 2007 puis 2016 ([9]), est un indice intercalibré à l'échelle européenne, intégré dans les arrêtés nationaux d'évaluation et utilisé en routine pour la surveillance écologique des cours d'eau de France métropolitaine [10] [11]. Son calcul est basé sur une liste de 838 taxons clés dits contributifs qui présentent un profil écologique rendant compte de leur probabilité de présence le long d'un gradient de 7 classes de qualité, construit à partir de paramètres physico-chimiques (pH, conductivité, saturation en oxygène, demande biologique en oxygène, ammonium, nitrates, phosphates) [9]. En utilisant les relevés d'abondance floristique annuels échantillonnés dans les différentes stations du réseau de contrôle de surveillance (RCS), l'IBD fournit une note comprise entre 1 et 20 qui, transformée en ratio de qualité écologique, renseigne sur l'état des masses d'eau « cours d'eau » et permet de prioriser les actions de gestion.

Cependant, la composition des communautés de diatomées présentes sur le territoire évolue, en lien avec les pressions évoquées précédemment, mais aussi en fonction des avancées de la taxonomie qui réévalue fréquemment la classification des espèces au sein de complexes ou de genres. Ces événements peuvent rendre en partie obsolète la liste de taxons clés et le profil écologique qui leur est associé. L'échantillonnage régulier des diatomées sur les 15 dernières années a également mis en évidence l'apparition d'espèces invasives dans les relevés. Certaines espèces possèdent donc un profil écologique qui n'est plus d'actualité au vu des modifications taxonomiques récentes, et d'autres espèces, dorénavant présentes en abondance significative, nécessitent l'association d'un profil.

L'objectif de ce rapport consiste donc à réviser/déterminer le profil écologique de certains taxons, nommés "taxons cibles". Dans ce but, les experts diatomistes des directions régionales de l'Office Français de la Biodiversité (OFB) ont dressé la liste de 71 taxons cibles pour lesquels il serait nécessaire de réévaluer ou de déterminer le profil écologique. Néanmoins, la condition de départ de notre travail consistait à ne pas (re)calculer de profil pour ces taxons, ce qui aurait demandé de recalculer le profil de tous les taxons clés de l'IBD. Cela aurait impliqué de générer une nouvelle version de l'indice susceptible d'impacter significativement l'évaluation actuelle de l'état écologique des masses d'eau, et qu'il aurait été nécessaire de ré-intercalibrer. Cette refonte complète n'étant pas envisageable pour l'instant, nous proposons ici une approche permettant de mettre à jour la liste des taxons clés, sans recalculer à proprement parler de nouveaux profils, et en s'appuyant sur l'avis d'expert.

¹ Directive 2000/60/ce du parlement européen et du conseil du 23 octobre 2000 établissant un cadre pour une politique communautaire dans le domaine de l'eau

2 MÉTHODE

2.1 ORIGINE DES DONNÉES EXPLOITÉES

La première étape de notre travail a consisté à récupérer les données issues du prélèvement des stations d'échantillonnage du réseau national de surveillance DCE, entre 2007 et 2021. Pour cela, deux bases ont été exploitées, la base NAIADES, base de données publique disponible à l'adresse : <https://naiades.eaufrance.fr> ; et la base PANDORE, base de requêtage SQL interne à INRAE. Dans chacune d'elles, les prélèvements de diatomées ont été renseignés annuellement, entre mai et septembre. La fusion des résultats obtenus à partir de ces deux bases nous a permis de construire une première banque de données contenant les relevés de 8 737 stations.

Au total, 55 664 listes floristiques associées à un code de station et une date comprise entre 2007 et 2021 ont été récupérées et compilées, ce qui correspond à 2783 taxons différents. Chaque relevé de la base se présente sous la forme: Date / Site / Espèce / Abondance relative. L'abondance en ‰ d'un taxon est déterminée pour chaque date et chaque site d'échantillonnage, relativement à tous les autres taxons prélevés à cette même date sur ce même site.

2.2 TAXONS CIBLES

2.2.1 Liste proposée par les experts diatomistes

Les experts diatomistes de l'OFB nous ont fait parvenir la liste des taxons nécessitant selon eux la révision de leur profil écologique, ou l'attribution d'un profil afin de les rendre contributifs au calcul de l'IBD. Parmi les 71 taxons référencés, 24 d'entre eux ont bénéficié d'une proposition de profil sur avis d'expert, proposition correspondant au profil écologique d'un taxon clé contribuant à l'IBD et ayant selon eux une écologie équivalente. En revanche, 47 taxons ne bénéficiaient pas d'une telle proposition. Le Tableau 1 présente les taxons bénéficiant d'une proposition de profil écologique équivalent, et le Tableau 2 présente les taxons sans profil écologique proposé.

Tableau 1: Taxon avec proposition de profil écologique

Taxon	Profil proposé	Codes 4 lettres
<i>Achnanthydium microcephalum</i>	<i>Achnanthydium saprophilum</i>	ADMC -> ADSA
<i>Achnanthydium delmontii</i>	<i>Achnanthydium minutissimum</i>	ADMO -> ADMI
<i>Achnanthydium subhudsonis</i> var. <i>kraeuselii</i>	<i>Achnanthydium subhudsonis</i>	ADSK -> ADSH
<i>Achnanthydium tropicocatenatum</i>	<i>Achnanthydium catenatum</i>	ADTC -> ADCT
<i>Achnanthydium hoffmannii</i>	<i>Achnanthydium minutissimum</i>	AHOF -> ADMI
<i>Achnanthydium zhakovschikovii</i>	<i>Achnanthydium pfisteri</i>	AZHA -> APFI
<i>Cymbella excisiformis</i>	<i>Cymbella affinis</i>	CEXF -> CAFF
<i>Cocconeis rouxii</i>	<i>Cocconeis pseudolineata</i>	CROU -> COPL
<i>Cymbella subhelvetica</i>	<i>Cymbella helvetica</i>	CSBH -> CHEL
<i>Eunotia arcubus</i> var. <i>arcubus</i>	<i>Eunotia arcus</i>	EARB -> EARC
<i>Encyonopsis alpina</i>	<i>Encyonopsis minuta</i>	ECAL -> ECPM
<i>Eunotia cantonati</i>	<i>Eunotia bilunaris</i>	ECTO -> EBLU
<i>Eunotia juettnerae</i>	<i>Eunotia bilunaris</i>	EJUE -> EBLU
<i>Eunotia pseudogroenlandica</i>	<i>Eunotia minor</i>	EPSG -> EMIN
<i>Fragilaria perdelicatissima</i>	<i>Ulnaria delicatissima</i>	FPDE -> UDEL
<i>Gomphonema minusculum</i>	<i>Gomphonema elegantissimum</i>	GMIS -> GELG
<i>Hippodonta pseudacceptata</i>	<i>Navicula gregaria</i>	HPDA -> NGRE
<i>Luticola hlubikovae</i>	<i>Luticola goeppertiana</i>	LHLU -> LGOP
<i>Nanofrustulum sopotensis</i>	<i>Stauronira venter</i>	NFSO -> SSVE
<i>Nitzschia soratensis</i>	<i>Nitzschia inconspicua</i>	NSTS -> NINC
<i>Platessa saxonica</i>	<i>Psammothidium helveticum</i>	PSXO -> PHEL
<i>Sellaphora bosniaca</i>	<i>Sellaphora difficillima</i>	SBOS -> SDIF
<i>Sellaphora crassulexigua</i>	<i>Sellaphora nigri</i>	SCRA -> SNIG
<i>Sellaphora capitata</i>	<i>Sellaphora pupula</i>	SECA -> SPUP

Tableau 2: Taxons sans proposition de profil écologique

Taxon	Codes 4 lettres	Taxon	Codes 4 lettres
<i>Achnantheidium atomoides</i>	ADAM	<i>Gomphonema cuneolus</i>	GCUN
<i>Achnantheidium caravelense</i>	ADCV	<i>Gomphonema innocens</i>	GINN
<i>Achnantheidium minutissimum</i>	ADMI	<i>Gomphonema parapygmaeum</i>	GPPY
<i>Achnantheidium druartii</i>	ADRU	<i>Gomphonema pumilum</i>	GPUM
<i>Adlafia baicalensis</i>	AFBA	<i>Gomphonema tenocultum</i>	GTNO
<i>Adlafia langebertalotii</i>	ALBL	<i>Gomphonema varioeduncum</i>	GVRD
<i>Amphora macedoniensis</i>	AMCD	<i>Hippodonta ruthnielseniae</i>	HRUT
<i>Cymbopleura pyrenaica</i>	CBPY	<i>Halamphora thumensis</i>	HTHU
<i>Cymbella lancettula</i>	CLTL	<i>Navicula difficillimoides</i>	NDDF
<i>Coconeis placentula</i> (actuel)	CPLA	<i>Navicula phylleptosoma</i>	NPHP
<i>Cymbella subtruncata</i> var. <i>subtruncata</i>	CSUT	<i>Navicula subalpina</i>	NSBN
<i>Cymbella tridentina</i>	CTDE	<i>Nitzschia costei</i>	NYCO
<i>Encyonema bonapartei</i>	EBNA	<i>Planothidium reichardtii</i>	PLRC
<i>Encyonopsis neerlandica</i>	ENEE	<i>Planothidium minutissimum</i>	PMNT
<i>Encyonema simile</i>	ENSI	<i>Punctastriata ovalis</i>	POVA
<i>Fragilaria candidagilae</i>	FCAD	<i>Pseudostaurosira polonica</i>	PSPO
<i>Fragilaria canariensis</i>	FCAN	<i>Sellaphora atomoides</i>	SEAT
<i>Fragilaria gracilis</i>	FGRA	<i>Sellaphora saprotolerans</i>	SESP
<i>Fragilaria microvaucheriae</i>	FMIV	<i>Sellaphora nigri</i>	SNIG
<i>Fragilaria nevadensis</i>	FNEV	<i>Sellaphora pseudoarvensis</i>	SPDV
<i>Fragilaria neointermedia</i>	FNIN	<i>Sellaphora raederae</i>	SRAE
<i>Fragilaria pararumpens</i>	FPRU	<i>staurosira robusta</i>	SRBU
<i>Gomphonema angustivalva</i>	GAGV	<i>Sellaphora schauburgii</i>	SSBG
<i>Gomphonema auritum</i>	GAUR		

Parmi les taxons cibles proposés par les experts diatomistes de l'OFB, certains nécessitent l'affectation d'un nouveau profil, en lien avec les récentes avancées taxonomiques ([13], [14]). Par exemple :

- *Achnantheidium microcephalum* (ADMC) possède à l'heure actuelle le profil écologique de *Achnantheidium minutissimum* (ADMI) constituant un complexe taxonomique. Cependant une partie des identifications notées ADMI correspondraient en réalité à ADCMC, qui nécessite donc l'affectation d'un profil distinct de celui d'ADMI.
- L'ancienne appellation *Eolimna minima sensu lato* (EOMI) qui regroupait en réalité SCRA, SEAT et SNIG n'est aujourd'hui plus d'actualité ([13]), et chaque taxon initialement présent dans le complexe doit bénéficier de son propre un profil écologique.
- La description de l'actuel *Cocconeis placentula* (depuis la publication de Jahn *et al.* en 2009 [14]) correspond à *Cocconeis euglyptoides* (CEUO). Ce « nouveau CPLA » ayant été créé, il serait nécessaire de lui associer un profil, différent de celui attribué à l'ancienne version de CPLA.

2.2.2 Transcodage

Etant donné l'évolution rapide de la taxonomie des diatomées, tous les opérateurs sur le territoire n'utilisent pas forcément la même appellation pour un même taxon. En résultent des listes taxonomiques parfois hétérogènes, bancarisées dans les bases nationales sans prétraitement. Pour s'assurer de travailler à partir des appellations taxonomiques les plus récentes (et éviter de considérer dans nos analyses comme deux taxons différents ce qui correspond uniquement à deux appellations différentes), nous avons utilisé une table de transcodage pour harmoniser la base. Elle considère pour chaque taxon son appellation la plus récente. Ce n'est qu'après avoir effectué ce transcodage sur toute la base de données que nous avons pu démarrer les analyses.

2.2.3 Base de données de travail

Des difficultés liées à l'harmonisation taxonomique ont été mises en évidence lors de l'exploration approfondie de la base initialement constituée, notamment la datation précise de l'intégration de changements d'appellation de certains taxons suite à une publication. Par exemple, la transition de *Eolimna minima sensu stricto* vers l'appellation *Sellaphora nigri* a été complexe étant donné que *Eolimna minima sensu stricto* appartenait à un groupe taxonomique (*Eolimna minima* dorénavant considéré *sensu lato*) comprenant *Sellaphora crassulexigua*, *Sellaphora nigri* et *Sellaphora atomoides* ([13]). La présence progressive de *Sellaphora nigri* dans les relevés empêche une détermination précise du moment de la transition, avec des relevés alternant entre *Eolimna minima sensus stricto* et *Sellaphora nigri* pour une même année. Pour s'affranchir au mieux de ces difficultés, nous avons analysé les données pour lesquelles nous sommes certains que la transition d'appellation a été réalisée pour tous les taxons cibles considérés, à savoir la période 2015-2021 qui correspond à 7788 stations et 25196 listes floristiques. Il s'agit de notre base de données de travail. La Figure 1 présente la répartition des stations d'échantillonnage sur le territoire pour la période concernée.



Figure 1: Répartition des sites échantillonnés entre 2015 et 2021.

Pour visualiser l'importance des taxons cibles dans les relevés floristiques, le nombre d'occurrences de chacun d'entre eux dans les relevés entre 2015 et 2021 est répertorié en annexe. Ce tableau présente une grande hétérogénéité dans les données, avec des occurrences allant de 2 à 22 556, ce qui correspond à des taxons présents dans 90 % des relevés contre d'autres présents dans seulement moins d'1 % (0.008) des relevés. Douze taxons (ADSK, CBPY, CROX, CTDE, ECTO, EJUE, ENEE, EPSG, GPPY, HRUT, PSPO et SBOS) présentent des occurrences inférieures à 30, ce qui est insuffisant pour appliquer des approches statistiques de façon fiable [15]. Nous avons donc décidé de s'affranchir de ces taxons dans les analyses ultérieures, et de se référer aux dires des experts diatomées pour leur affilier un profil écologique, dans la mesure où ils les considèrent essentiels dans la mise à niveau de l'IBD.

2.3 DÉVELOPPEMENT D'UNE MÉTHODE POUR ASSOCIER DES PROFILS ÉCOLOGIQUES AUX TAXONS CIBLES

Afin d'être le plus clair possible dans la démarche utilisée, nous proposons en Figure 2 un schéma qui récapitule le processus général d'association d'un profil écologique à un taxon cible, processus décrit en détail ensuite.

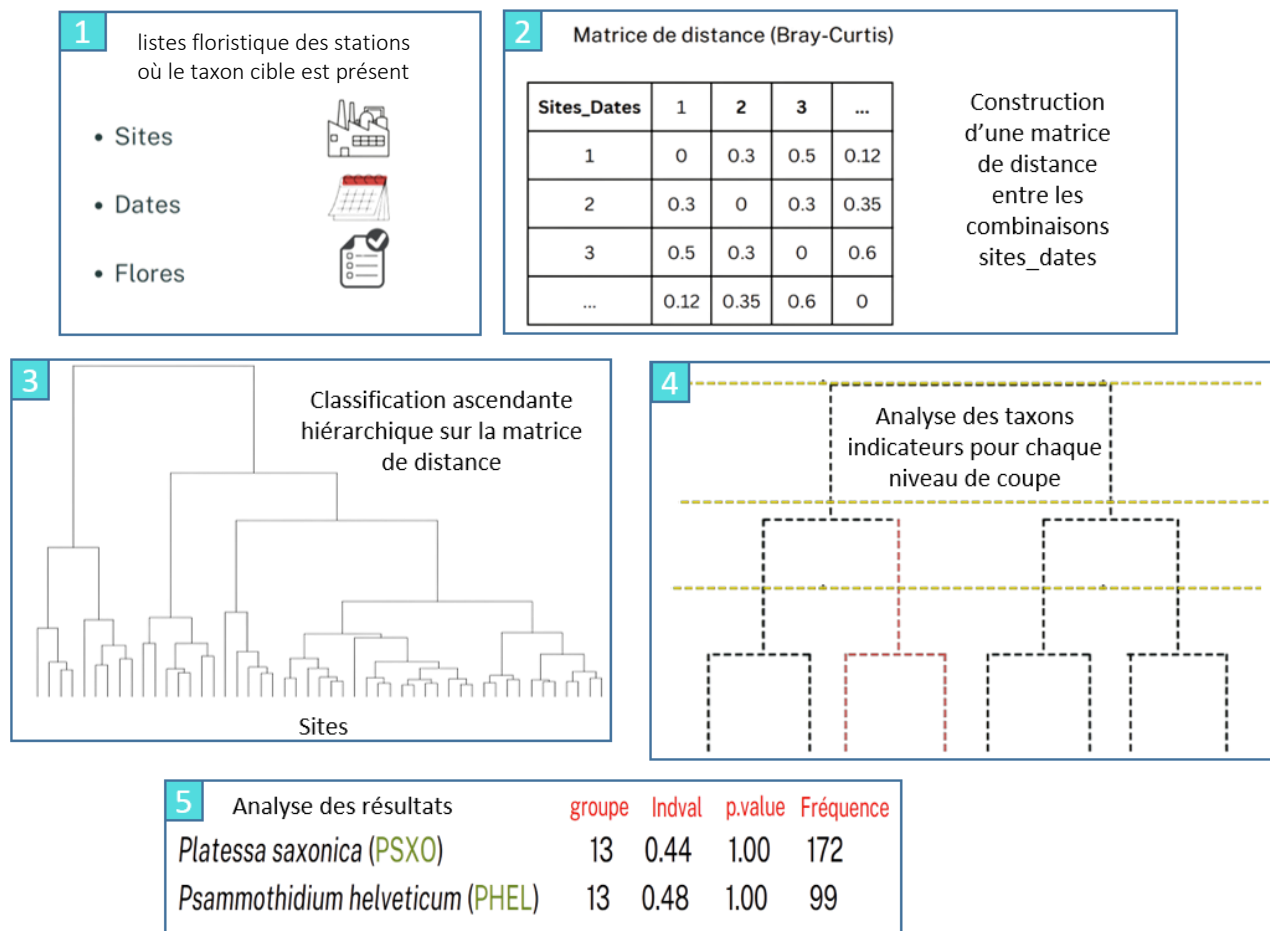


Figure 2: Schéma de la méthode utilisée pour associer un profil écologique aux taxons cibles

Nous rappelons ici que l'objectif méthodologique n'est pas de calculer de nouveaux profils écologiques pour les taxons cibles, ce qui entraînerait une refonte de l'indice, mais plutôt de leur associer des profils préexistants parmi ceux des taxons clés permettant le calcul de l'IBD. Développer une telle méthode implique d'être capable de comparer la distribution écologique des différents taxons de notre base de données, et de repérer ceux qui semblent co-occurrencer, c'est-à-dire ceux qui semblent présenter des *preferenda* écologiques équivalents. Notre approche se base sur la mise en œuvre de classifications ascendantes hiérarchiques ([16]) associées à la détermination de taxons indicateurs des groupes obtenus, *via* la méthode de l'Indicator species analysis ([17]). Notre démarche consiste également à dialoguer de façon étroite avec l'expertise, qui vient *a posteriori* confirmer ou compléter les résultats lorsque cela s'avère nécessaire.

2.3.1 Construction d'une base de données de comparaison

La première étape consiste donc à déterminer une méthode fiable de comparaison de la distribution des taxons au sein de la base de données (restreinte à la période 2015-2021, cf paragraphe 2.2.3). Il est important de considérer la globalité des taxons des relevés floristiques, y compris ceux qui n'ont pas de profil écologique, dans ces comparaisons avec nos taxons cibles. Inclure uniquement les taxons clés de l'indice dans cette phase de comparaison ne prendrait pas en compte l'ensemble des données, et ne permettrait pas d'aboutir à des associations pertinentes.

Le format de nos données permet d'extraire facilement toutes les informations relatives à un taxon cible. Pour comparer la distribution de ce taxon avec celle des autres, nous avons commencé par réduire la base aux seuls sites et dates où ce taxon est observé. Cela permet de créer un sous-jeu de données contenant uniquement des listes floristiques dans lesquelles le taxon cible est présent. Cette étape nous permet de s'affranchir de la variabilité apportée par le reste des relevés, et de nous focaliser sur le taxon cible et les autres taxons qui s'y associent de façon significative dans la base (site + date). Un taxon pouvant être observé sur un même site sur plusieurs années consécutives, nous avons identifié de manière unique chaque prélèvement du sous-jeu de données par sa date et son numéro de station. Cela nous assure de prendre en compte la variabilité des communautés diatomiques présentes sur une même station à différentes années.

Cette nouvelle configuration nous permet d'extraire et de comparer les compositions floristiques de chaque combinaison Site_Date. Pour cela, nous avons commencé par standardiser les abondances relatives par la méthode de Hellinger [17]. Cette étape nous permet de standardiser la distribution des taxons sur une échelle commune d'abondance, ce qui facilite les comparaisons et évite la sous- ou la sur-représentation de certains. Nous avons ensuite utilisé la distance de Bray-Curtis [18] pour évaluer entre 0 et 1 la dissimilarité floristique entre chaque combinaison Site_Date. Le Tableau 3 présente un exemple de cette matrice de distance entre relevés. Plus cette distance est proche de 1, plus les relevés sont différents du point de vue de leur composition floristique.

Tableau 3: Extrait d'une matrice de distance de Bray-Curtis

	02110050_2020_07_09	02061970_2020_07_010	02045200_2020_07_03
02045200_2020_07_03	0.8320558		
02061970_2020_07_010	0.9031096	0.8913325	
02110050_2020_07_09	0.8478347	0.6696994	0.9501146

À partir de ces matrices de dissimilarité (3) issues de chaque sous-jeu de données, il est possible d'appliquer une classification ascendante hiérarchique (CAH) [16] pour construire un dendrogramme sur lequel chaque combinaison Site_Date est positionnée relativement aux autres en fonction de la distance de Bray-Curtis (Figure 3). Les relevés sont donc regroupés en fonction de la similarité de leur composition floristique. C'est sur la base de ce dendrogramme que nous avons construit notre méthode de comparaison de la distribution des taxons.

2.3.2 Taxons indicateurs selon la méthode IndVal

La détermination de taxons indicateurs, autrement appelée analyse « IndVal » et initialement définie par Dufrêne & Legendre [17], est une méthode qui permet d'évaluer l'affinité des taxons avec différents groupes de sites ou de conditions environnementales. Elle est couramment utilisée à la suite d'une CAH pour identifier les taxons indicateurs de chaque groupe de relevés, c'est-à-dire les taxons fortement associés à des conditions spécifiques. Cette analyse repose sur deux paramètres principaux, la spécificité et la fidélité.

La spécificité (souvent notée A) d'un taxon est évaluée en comparant sa fréquence ou son abondance moyenne dans un groupe de sites donné par rapport à l'ensemble des sites échantillonnés. Plus elle est élevée, plus le taxon est associé à ce groupe de sites spécifique. La fidélité (souvent notée B) représente le nombre de relevés occupés par le taxon parmi les relevés d'un groupe. La spécificité est de 100% quand le taxon n'occupe qu'un groupe et la fidélité est de 100% lorsque le taxon est présent dans tous les relevés d'un groupe.

La valeur d'IndVal est ensuite calculée de la manière suivante :

$$IndVal = A * B * 100$$

Après avoir effectué l'analyse IndVal, des tests de permutation déterminent si l'association d'un taxon avec un groupe de sites est statistiquement significative. Ces tests de permutation permettent de quantifier la probabilité d'observer une association aussi forte, voire plus forte, entre le taxon et le groupe de sites par simple hasard.

On effectue ainsi des permutations aléatoires de tous les taxons entre les groupes tout en conservant les totaux d'occurrence et d'abondance de chaque taxon. Pour chaque permutation, on calcule l'indice du taxon cible pour le groupe d'intérêt. En répétant ce processus 500 fois, on crée alors une distribution des valeurs d'IndVal du taxon considéré basée uniquement sur le hasard.

En comparant l'association observée dans les données réelles à cette distribution, on peut calculer la probabilité p d'obtenir une association aussi forte ou plus forte que celle observée par simple hasard. Si cette valeur de p est inférieure à un seuil prédéfini (nous utilisons un seuil de 0,05), on considère alors que l'association entre le taxon et le groupe de sites est statistiquement significative.

2.3.3 Développement de l'approche IndVal pour identifier les taxons aux distributions équivalentes

Il est possible en scindant le dendrogramme à un niveau de coupe donné de créer des groupes caractérisés par des communautés homogènes de diatomées et donc typiques de conditions environnementales particulières. Néanmoins, la complexité élevée des données de relevés diatomiques rend difficile la définition d'une hauteur de coupe précise du dendrogramme. De ce fait, nous avons dû développer une approche particulière.

A partir du dendrogramme, nous avons effectué une coupe pour chaque hauteur de nœud (Figure 3). Pour un taxon échantillonné à N reprises, cela revient à découper le dendrogramme à $N-1$ hauteurs différentes. Pour chaque niveau de coupe, des niveaux différents d'agrégation des relevés en groupes homogènes sont obtenus. Nous avons réalisé l'analyse IndVal pour chaque niveau de coupe puis récupéré les valeurs obtenues pour chaque taxon cible lorsque celui-ci était considéré indicateur d'un groupe. Nous avons ensuite généré un tableau de synthèse (non présenté ici) regroupant les informations relatives aux taxons indicateurs des mêmes groupes que nos taxons cibles. Dans ce tableau, nous avons déterminé pour chaque taxon sa valeur moyenne d'IndVal, sa spécificité et fidélité moyennes, son abondance médiane et le nombre de fois qu'il a été observé indicateur en même temps que le taxon cible.

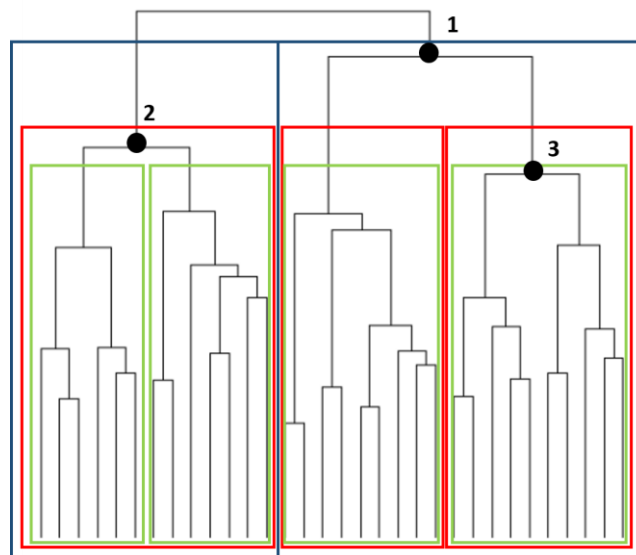


Figure 3: Exemple des groupes générés pour 3 hauteurs de coupes différentes du dendrogramme d'ADTC. Les points noirs représentent les nœuds et les rectangles de couleurs les groupes générés, en bleu pour le nœud numéro 1, en rouge pour le nœud numéro 2 et en vert pour le nœud numéro 3.

2.3.4 Méthodologie générale d'attribution d'un profil écologique

Grâce à la méthode IndVal, les taxons indicateurs des groupes générés sont déterminés. Ces taxons indicateurs présentent des distributions écologiques comparables. Parmi eux, certains sont des taxons clés de l'IBD, et ont donc un profil écologique déterminé. Lorsqu'un taxon cible est sélectionné comme taxon indicateur, il est donc associé à d'autres taxons indicateurs, possédant déjà pour certains un profil écologique. Ce profil écologique peut donc être proposé comme candidat pour le taxon cible.

Pour les espèces vues moins de 30 fois (et donc écartées des analyses statistiques) :

- Les taxons bénéficiant d'une proposition de profil écologique de la part des experts diatomistes (voir Tableau 1) reçoivent ce profil.
- Ceux ne bénéficiant pas d'un tel avis d'expert ne pourront pas être considérés comme taxon clé et ne contribueront pas au calcul de l'IBD. Leur cas pourra être réévalué ultérieurement lorsque davantage de données relatives à leur observation seront disponibles.

Pour les autres taxons (observés plus de 30 fois dans notre base de données) :

Il est très rare de n'observer qu'un ou deux taxons indicateurs d'un groupe donné, en particulier lorsque l'on travaille avec des données complexes. Dans notre cas, un taxon cible est le plus souvent associé à plusieurs taxons. Comme il n'existe pas de critère défini permettant de sélectionner le "meilleur" candidat, nous avons procédé de la manière suivante:

- Les taxons cibles retrouvés indicateurs d'un groupe avec le taxon proposé par les experts diatomées se sont vus attribuer le profil correspondant. Par exemple, d'après le Tableau 2, si ADSA figure avec ADMC de façon significative parmi les taxons indicateurs du dendrogramme d'ADMC, la proposition des experts est validée et le profil d'ADSA est attribué à ADMC.

Dans le cas inverse où le taxon proposé par les experts ne se retrouverait pas associé au taxon cible parmi les indicateurs, nous proposons aux experts un taxon indicateur du même genre et présentant un profil écologique similaire à celui initialement proposé. Nous avons conforté cette proposition en vérifiant la cohérence des *preferenda* écologiques de ces deux taxons (cible et indicateur) à partir de la publication de Carayon *et al.* [19].

- Pour les taxons cibles sans proposition de profil de la part des experts, nous avons recherché parmi les taxons indicateurs associés ceux du même genre possédant un profil écologique (selon l'IBD) et présentant des *preferenda* écologiques (selon Carayon *et al.*) comparables.
- Dans les rares cas où les *preferenda* écologiques du taxon cible n'étaient pas déterminés par Carayon *et al.*, [19] nous avons sélectionné les taxons indicateurs du même genre.

Dans chaque cas, nos propositions ont été soumises à l'avis des experts, dans un dialogue continu permettant si besoin des réajustements quant au taxon candidat le plus approprié.

2.4 IMPACT DE L'INTÉGRATION DES TAXONS CIBLES SUR L'ÉVALUATION DE L'ÉTAT ÉCOLOGIQUE

Afin d'évaluer la répercussion de la mise à jour des profils indiciaires sur les notations IBD des masses d'eau, nous avons calculé les notes IBD d'un même jeu de données (intégralité des relevés présents sur la base NAIADES, 2007-2021) en incluant d'abord l'ancienne liste des taxons cibles puis la nouvelle. Les algorithmes de calcul de l'IBD que nous avons utilisés sont ceux disponibles sur le site du Système d'évaluation de l'état des eaux (SEEE): <https://seee.eaufrance.fr/>. La méthodologie de calcul de base de l'indice biologique diatomées peut être retrouvé dans la publication de Coste *et al.* [9]

Nous avons ensuite comparé statistiquement les distributions des notes des deux indices par le biais de test de Wilcoxon pour données appariées, après avoir rejeté l'hypothèse de normalité de la distribution des notes. Dans un second temps, nous avons analysé la corrélation entre les résultats obtenus en calculant l'indice basé sur la liste initiale de taxons clés (IBD1) et ceux obtenus à partir de l'indice contenant la nouvelle liste (IBD2) au moyen d'un test de corrélation de Spearman (non-paramétrique). Nous avons enfin évalué la corrélation entre la différence de résultats entre l'IBD1 et l'IBD2 (IBD1-IBD2) et les notes fournies par l'IBD2 au moyen de tests de corrélation de Spearman, pour étudier le sens de la variabilité éventuelle de notation entre l'IBD2 et l'IBD1.

2.5 DÉVELOPPEMENT D'UNE APPLICATION DE VISUALISATION DE LA BASE NAIADES

Les données diatomées sont des données complexes qui nécessitent des connaissances approfondies sur les distributions géographiques et les préférences écologiques des taxons. En ce sens, NAIADES se veut être un outil qui permet un accès rapide aux données floristiques sur les diatomées à l'échelle de la France. Le potentiel offert par cette banque de données est important mais encore sous-exploité. Fournir un outil capable de valoriser graphiquement les données NAIADES faciliterait leur appropriation et leur utilisation afin de mieux connaître notamment la répartition géographique des taxons et leurs co-occurrences.

L'application Visual Diatoms que nous avons développée en appui au travail présenté ici est une application web multilingue qui communique directement avec la base NAIADES (plus précisément sa dernière mise à jour), et propose un accès complet et formaté aux données diatomées qu'elle contient. Un aperçu de la page d'accueil est disponible en Figure 4.



Figure 4: Page d'accueil de l'application Visual Diatoms

Deux onglets principaux disponibles à la sélection permettent à l'utilisateur d'explorer la base de deux façons, soit directement *via* l'onglet "Données Brutes" qui répertorie tous les prélèvements diatomés disponibles dans NAIADES (Figure 5), soit par l'onglet "Visualisation" qui permet de concentrer l'observation à l'échelle des taxons (Figure 6). Ce dernier offre la possibilité à l'utilisateur d'analyser la répartition temporelle et spatiale d'un ou plusieurs taxons, de comparer leurs préférences écologiques, et de télécharger les données formatées du ou des taxons ciblés directement sur ordinateur.

	Date	Station	commune	Longitude	Latitude	Taxon	Abondance relative (%)
1	2007-07-02	2115200	Longjumeau	5.8	49.54	Achnanthyrium eutrophilum (ADEU)	2.5
2	2007-07-02	2115200	Longjumeau	5.8	49.54	Achnanthyrium lineare (ACL)	2.5
3	2007-07-02	2115200	Longjumeau	5.8	49.54	Achnanthyrium minutissimum (ADMI)	37.5
4	2007-07-02	2115200	Longjumeau	5.8	49.54	Achnanthyrium saprophilum (ADSA)	17.5
5	2007-07-02	2115200	Longjumeau	5.8	49.54	Amphora pediculus (APED)	110
6	2007-07-02	2115200	Longjumeau	5.8	49.54	Halamphora veneta (HVEN)	7.5
7	2007-07-02	2115200	Longjumeau	5.8	49.54	Cocconeis pediculus (CPED)	105
8	2007-07-02	2115200	Longjumeau	5.8	49.54	Cocconeis placentula var. euglypta (CPLE)	17.5
9	2007-07-02	2115200	Longjumeau	5.8	49.54	Craticula accomoda (CRAC)	2.5
10	2007-07-02	2115200	Longjumeau	5.8	49.54	Cyclotella meneghiniana (CMEN)	15

Figure 5: Onglet "Données Brutes" de l'application Visual Diatoms

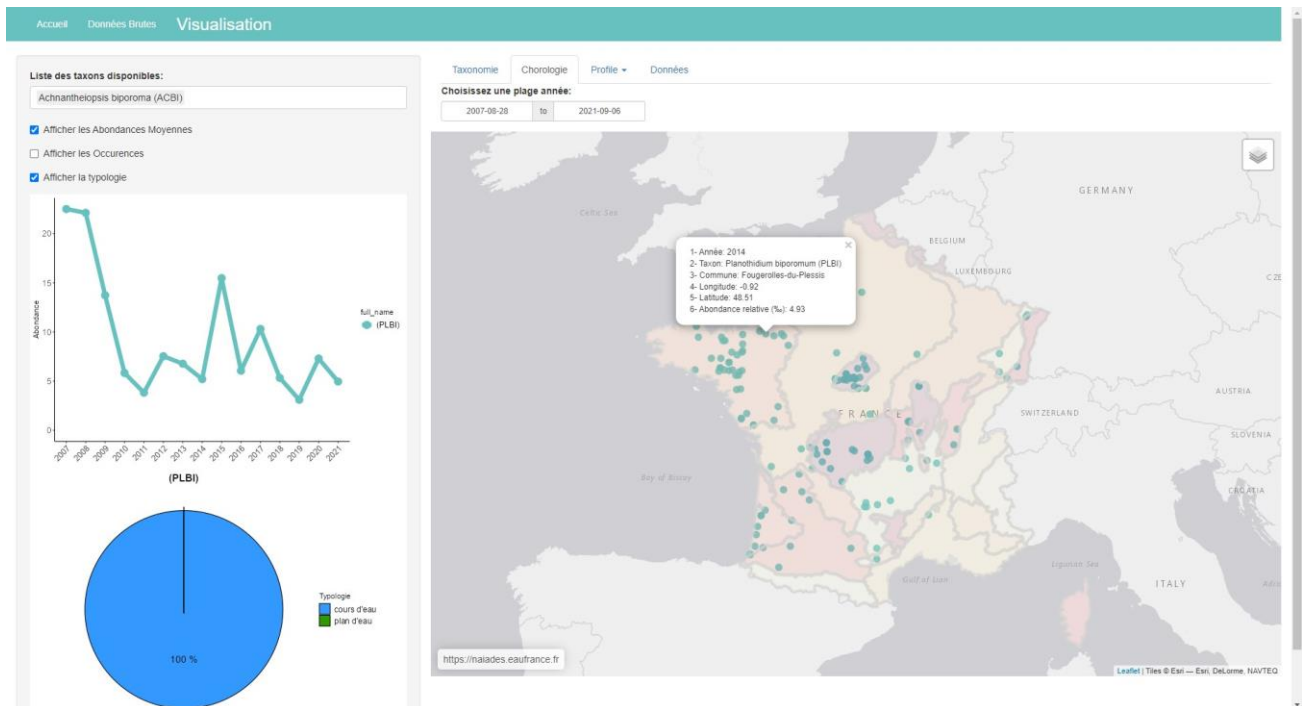


Figure 6: Onglet "Visualisation" de l'application Visual Diatoms

Avec cette application, nous espérons faciliter l'accès et la communication autour des données diatomées pour tous ceux qui sont amenés à les utiliser. Une version démo de l'application est disponible à l'adresse suivante :

<https://shiny.sk8.inrae.fr/app/aqua-eabx-visual-diatoms>

3 RÉSULTATS / DISCUSSION

3.1 BASE DE DONNÉES DE TRAVAIL

La base de données finale que nous avons utilisée pour conduire les analyses (période 2015-2021) est décrite dans cette section. La Figure 7 illustre dans un premier temps le contenu de la base.

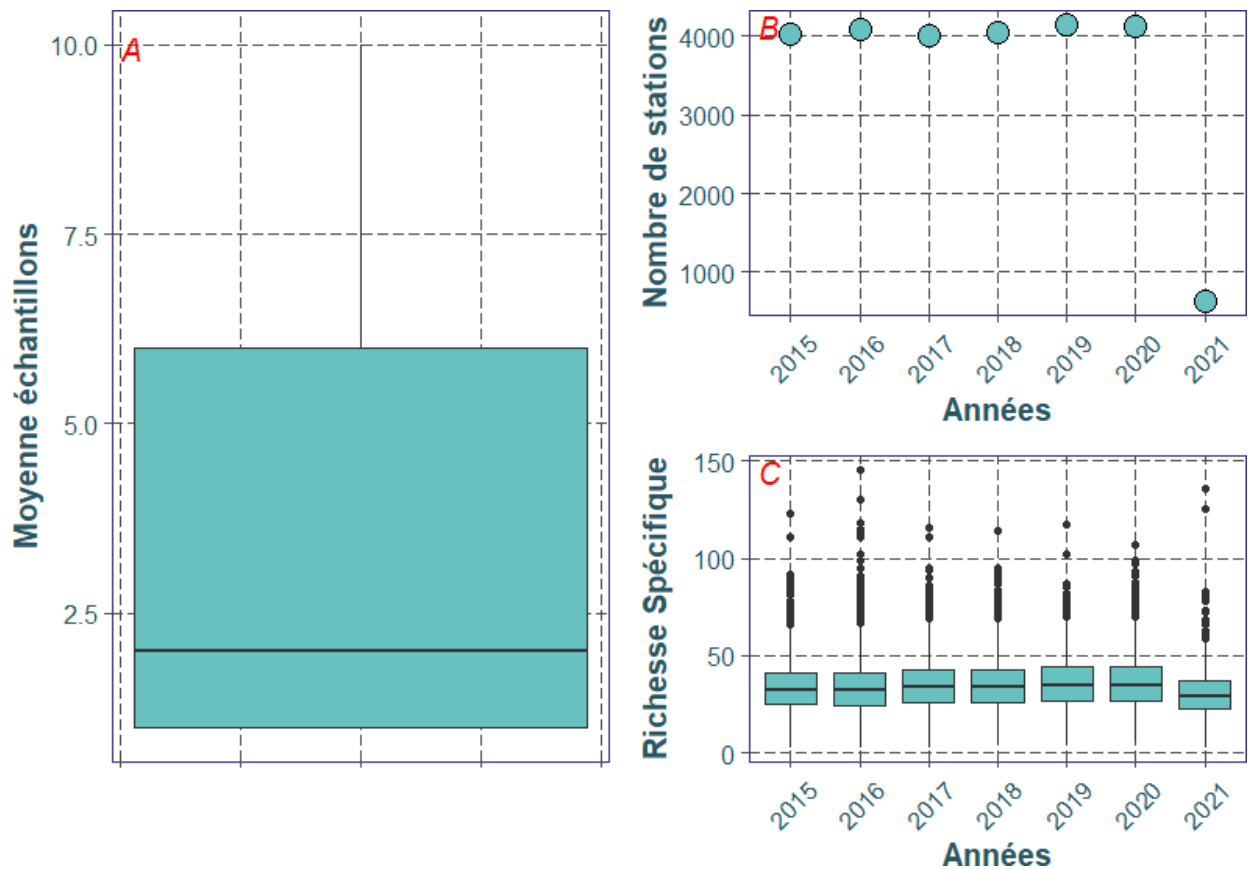


Figure 7: Représentations graphiques du contenu de la base de données utilisée pour les analyses : A, box-plot du nombre d'échantillons par station, B, Evolution du nombre de relevés disponibles en fonction des années et C, Evolution de la richesse spécifique des stations en fonction des années.

D'après la Figure 7, le nombre de stations échantillonnées chaque année est constant. L'année 2021 présente un très faible échantillonnage qui s'explique par le délai de mise à jour de la base NAIADES qui peut parfois être long d'une année sur l'autre. Toutes les données de l'année 2021 n'étaient pas disponibles au moment de notre étude. La richesse spécifique des relevés (Figure 7 C) est comparable d'une année sur l'autre et le nombre d'échantillons par station ne dépasse généralement pas 6, avec une médiane autour de 2 (Figure 7 B).

En ce sens, nous avons étudié en Figure 8 la répartition des stations d'échantillonnages d'après les hydro-écorégions (HER) de niveau 1 définies comme premier critère pour la typologie des rivières ([12]). À la suite de ce travail, nous avons estimé l'effort d'échantillonnage pour chacune des hydro-écorégions (Figure 9 et Tableau 4).

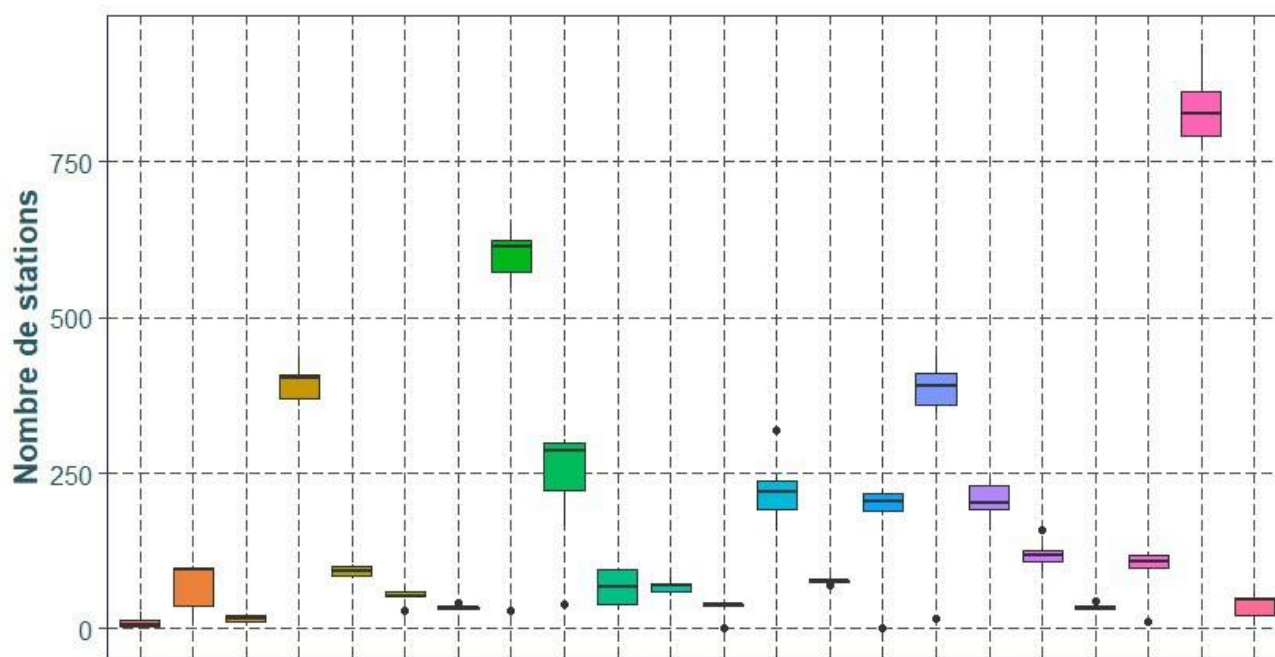


Figure 8: Figure représentant le nombre de stations échantillonnées par hydro-écorégions pour la période 2015- 2021

D'après la Figure 8, le nombre de stations échantillonnées sur la période 2015-2021 est assez variable d'une hydro-écorégion à l'autre avec des valeurs moyennes allant d'une dizaine à plus de 750. Cette forte variabilité peut s'expliquer par les surfaces respectives des HER et leur topographie : la difficulté d'accès de certaines régions comme les régions montagneuses implique un maillage de stations plus réduit par rapport à des régions moins vallonnées et donc plus faciles d'accès. Ces inégalités nous ont conduit à estimer un effort d'échantillonnage pour chaque hydro-écorégion (Tableau 4), en rapportant le nombre de stations échantillonnées à la surface de l'HER considérée.

Tableau 4: Effort d'échantillonnage par hydro-écorégion

Hydro-écorégion	Surface (km ²)	Nombre de stations	Effort échantillonnage
ALPES INTERNES	13736.810	54	0.33
CORSE	8681.129	51	0.50
LANDES	13263.846	83	0.60
PRÉALPES DU SUD	12144.670	77	0.50
CAUSSES AQUITAINS	10795.574	116	1.10
GRANDS CAUSSES	4410.842	49	1.11
CÉVENNES	8337.808	99	1.10
DÉPOTS ARGILO SABLEUX	12565.516	153	1.05
PYRÉNÉES	13531.037	168	1.05
VOSGES	7480.709	93	1.00
MASSIF CENTRAL NORD	30473.338	451	1.25
COTES CALCAIRES EST	46069.624	700	1.31
MÉDITERRANÉEN	32574.095	499	1.25
TABLES CALCAIRES	141040.684	2161	1.41
ARMORICAIN	61594.734	966	1.41
COTEAUX AQUITAINS	44211.780	797	1.73
ARDENNES	1625.594	32	1.78
MASSIF CENTRAL SUD	37066.375	732	1.83
PLAINE SAONE	9716.713	214	2.20
JURA-PRÉALPES DU NORD	24933.144	599	2.25
DÉPRESSIONS SÉDIMENTAIRES	7413.244	204	2.16
ALSACE	5764.001	192	2.80

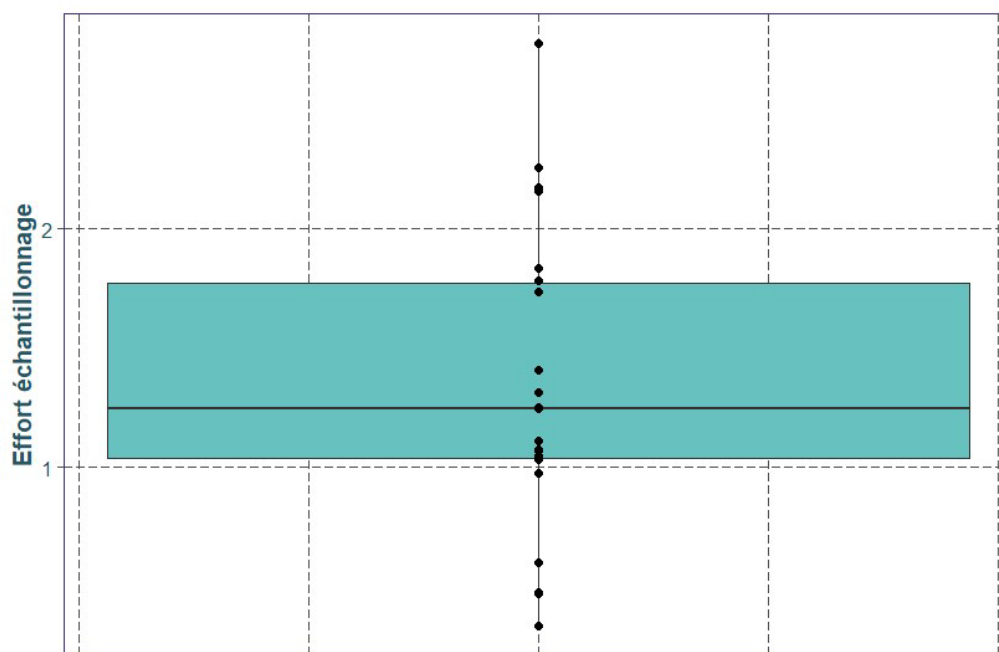


Figure 9: Box-plot de l'effort d'échantillonnage par hydro-écorégion

Le Tableau 4 indique que l'effort d'échantillonnage est compris entre 0.33 et 2.80. Malgré des valeurs extrêmes bien distinctes, 14 des 22 hydro-écorégions présentent un effort d'échantillonnage comparable. La Figure 9 montre les quelques outliers présents mais indique un effort d'échantillonnage global assez uniforme entre les HER.

3.2 ATTRIBUTION D'UN PROFIL ÉCOLOGIQUE AUX TAXONS CIBLES

Suite aux analyses et à partir des règles énoncées plus haut, nous avons dressé ci-dessous la liste récapitulative des décisions concernant les 71 taxons cibles (Tableau 5). Cette liste indique pour chacun s'il est retenu pour contribuer au calcul de l'IBD en tant que taxon clé et, le cas échéant, le profil écologique qui lui est associé.

Tableau 5: Profils attribués aux taxons cibles

Taxon cible	Profil attribué	Remarque
<i>Achnanthydium atomoides</i> (ADAM)	<i>Achnanthydium Eutrophilum</i> (ADEU)	
<i>Achnanthydium caravelense</i> (ADCV)	<i>Achnanthydium lineare</i> (ACLI)	
<i>Achnanthydium delmontii</i> (ADMO)	<i>Nitzschia dissipata</i> (NDIS)	
<i>Achnanthydium druartii</i> (ADRU)	<i>Achnanthydium straubianum</i> (ADSB)	
<i>Achnanthydium hoffmannii</i> (AHOF)	<i>Achnanthydium minutissima</i> (ADMI)	
<i>Achnanthydium microcephalum</i> (ADMC)	<i>Achnanthydium minutissimum</i> (ADMI)	
<i>Achnanthydium subhudsonis</i> var. <i>kraeuselii</i> (ADSK)	<i>Achnanthydium subhudsonis</i> (ADSH)	sur avis d'expert
<i>Achnanthydium tropicocatenatum</i> (ADTC)	<i>Achnanthydium catenatum</i> (ADCT)	
<i>Achnanthydium zhakovschikovii</i> (AZHA)	<i>Achnanthydium pfisteri</i> (APFI)	sur avis d'expert
<i>Adlafia baicalensis</i> (AFBA)		occurrence trop faible
<i>Adlafia langebertalotii</i> (ALBL)	<i>Adlafia minuscula</i> var. <i>minuscula</i> (ADMS)	sur avis d'expert
<i>Amphora macedoniensis</i> (AMCD)	<i>Amphora copulata</i> (ACOP)	
<i>Cocconeis placentula</i> (sensu Jahn)		en attente nouveaux codes Sandre
<i>Cocconeis rouxii</i> (CROX)		occurrence trop faible
<i>Cymbella excisiformis</i> (CEXF)	<i>Cymbella elvetica</i> (CHEL)	
<i>Cymbella lancettula</i> (CLTL)		occurrence trop faible
<i>Cymbella subhelvetica</i> (CSBH)	<i>Cymbella elvetica</i> (CHEL)	
<i>Cymbella subtruncata</i> var. <i>subtruncata</i> (CSUT)	<i>Cymbella elvetica</i> (CHEL)	
<i>Cymbella tridentina</i> (CTDE)	<i>Cymbella affinis</i> var. <i>affinis</i> (CAFF)	sur avis d'expert
<i>Cymbopleura pyrenaica</i> (CBPY)		occurrence trop faible
<i>Encyonema bonapartei</i> (EBNA)	<i>Encyonopsis subminuta</i> (ESUM)	
<i>Encyonema simile</i> (ENSI)	<i>Encyonema silesiacum</i> (ESLE)	
<i>Encyonopsis alpina</i> (ECAL)	<i>Encyonopsis minuta</i> (ECPM)	
<i>Encyonopsis neerlandica</i> (ENEE)	<i>Encyonopsis subminuta</i> (ESUM)	
<i>Eunotia arcubus</i> var. <i>arcubus</i> (EARB)	<i>Eunotia arcus</i> (EARC)	
<i>Eunotia cantonati</i> (ECTO)	<i>Eunotia bilunaris</i> (EBLU)	sur avis d'expert
<i>Eunotia juettnerae</i> (EJUE)	<i>Eunotia bilunaris</i> (EBLU)	sur avis d'expert
<i>Eunotia pseudogroenlandica</i> (EPSG)		occurrence trop faible
<i>Fragilaria candidagilae</i> (FCAD)	<i>Fragilaria vaucheriae</i> (FVAU)	
<i>Fragilaria gracilis</i> (FGRA)		en attente nouveaux codes Sandre
<i>Fragilaria microvaucheriae</i> (FMIV)	<i>Fragilaria vaucheriae</i> (FVAU)	
<i>Fragilaria neointermedia</i> (FNIN)	<i>Fragilaria mesolepta</i> (FMES)	
<i>Fragilaria nevadensis</i> (FNEV)	<i>Frustulia saxonica</i> (FSAX)	
<i>Fragilaria pararumpens</i> (FPRU)	<i>Fragilaria bidens</i> (FBID)	sur avis d'expert
<i>Fragilaria perdelicatissima</i> (FPDE)	<i>Ulnaria delicatissima</i> (UDEL)	

<i>Gomphonema angustivalva</i> (GAGV)	<i>Gomphonema exilissimum</i> (GEXL)	
<i>Gomphonema auritum</i> (GAUR)	<i>Gomphonema coronatum</i> (GCOR)	
<i>Gomphonema cuneolus</i> (GCUN)	<i>Gomphonema pumilum var. rigidum</i> (GPRI)	sur avis d'expert
<i>Gomphonema innocens</i> (GINN)		occurrence trop faible
<i>Gomphonema minusculum</i> (GMIS)	<i>Gomphonema elegantissimum</i> (GELG)	
<i>Gomphonema parapygmaeum</i> (GPPY)		non bancarisé à ce jour (cde Sandre depuis 2022)
<i>Gomphonema pumilum</i> (GPUM)		en attente nouveaux codes Sandre
<i>Gomphonema tenocultum</i> (GTNO)	<i>Gomphonema lateripunctatum</i> (GLAT)	sur avis d'expert
<i>Gomphonema varioreduncum</i> (GVRD)	<i>Gomphonema rhombicum</i> (GRHB)	
<i>Halamphora thumensis</i> (HTHU)	<i>Halamphora montana</i> (HLMO)	
<i>Hippodonta pseudacceptata</i> (HPDA)	<i>Navicula gregaria</i> (NGRE)	
<i>Hippodonta ruthnielseniae</i> (HRUT)		occurrence trop faible
<i>Luticola hlubikovae</i> (LHLU)	<i>Luticola goeppertiana</i> (LGOP)	
<i>Nanofrustulum sopotensis</i> (NFSO)	<i>Staurosira venter</i> (SSVE)	
<i>Navicula difficillimoides</i> (NDDF)	<i>Navicula germainii</i> (NGER)	
<i>Navicula phylleptosoma</i> (NPHP)		occurrence trop faible
<i>Navicula subalpina</i> (NSBN)		occurrence trop faible
<i>Nitzschia costei</i> (NYCO)	<i>Nitzschia fonticola</i> (NFON)	
<i>Nitzschia soratensis</i> (NSTS)	<i>Nitzschia fonticola</i> (NFON)	
<i>Planothidium minutissimum</i> (PMNT)	<i>Planothidium lanceolatum</i> (PTLA)	
<i>Planothidium reichardtii</i> (PLRC)	<i>Planothidium frequentissimum</i> (PLFR)	sur avis d'expert
<i>Platessa saxonica</i> (PSXO)	<i>Platessa oblongella</i> (POBL)	
<i>Pseudostaurosira polonica</i> (PSPO)		occurrence trop faible
<i>Pseudostaurosira robusta</i> (SRBU)	<i>Fragilaria crotonensis</i> (FCRO)	
<i>Punctastriata ovalis</i> (POVA)	<i>Punctastriata lancettula</i> (PULA)	
<i>Sellaphora atomoides</i> (SEAT)	<i>Sellaphora bacillum</i> (SEBA)	
<i>Sellaphora bosniaca</i> (SBOS)	<i>Sellaphora difficillima</i> (SDIF)	
<i>Sellaphora capitata</i> (SECA)	<i>Sellaphora pupula</i> (SPUP)	
<i>Sellaphora crassulexigua</i> (SCRA)	<i>Nitzschia amphibia</i> (NAMP)	
<i>Sellaphora nigri</i> (SNIG)	<i>Sellaphora nigri</i> (SNIG)	conserve son profil actuel
<i>Sellaphora pseudoarvensis</i> (SPDV)	<i>Sellaphora Rhombelliptica</i> (SRHE)	sur avis d'expert
<i>Sellaphora raederae</i> (SRAE)	<i>Navicula exilis</i> (NEXI)	
<i>Sellaphora saprotolerans</i> (SESP)	<i>Sellaphora pupula</i> (SPUP)	
<i>Sellaphora schauburgii</i> (SSBG)	<i>Sellaphora subrotundata</i> (SSRT)	sur avis d'expert
<i>Staurosirella canariensis</i> (SCAN)		occurrence trop faible

Dans cette nouvelle liste, 56 des 71 taxons cibles initiaux ont pu être caractérisés par un profil écologique.

3.3 IMPACT DES NOUVEAUX TAXONS CONTRIBUTIFS SUR LES NOTES D'IBD

Afin d'analyser l'impact de la modification de la liste de taxons contributifs de l'IBD sur la notation des masses d'eau, nous avons conduit des analyses statistiques sur les résultats fournis par les deux versions différentes d'IBD, à savoir ceux fournis par l'IBD calculé à partir de l'ancienne liste (noté IBD1) et ceux fournis par l'IBD calculé à partir de la nouvelle liste (noté IBD2).

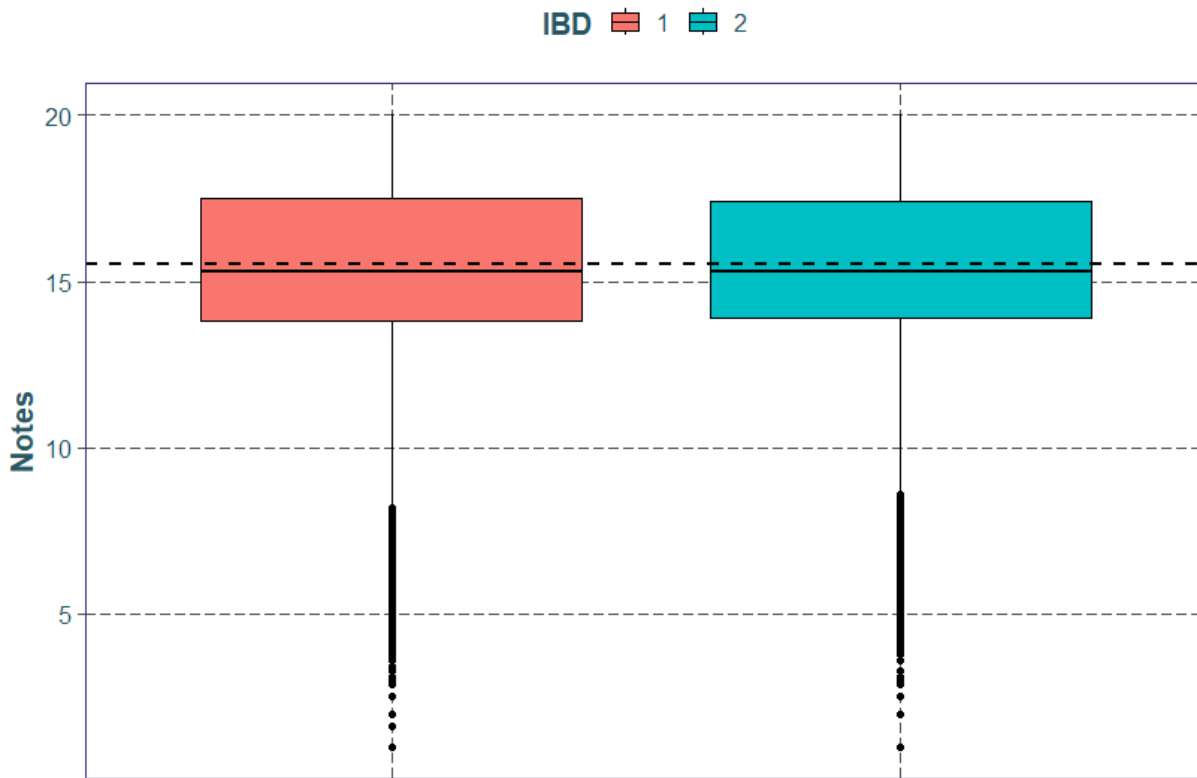


Figure 10: Box-plot de la distribution des résultats pour l'IBD calculé avec l'ancienne (IBD1) et la nouvelle (IBD2) liste de taxons clés

La Figure 10 montre une très légère différence entre les notes fournies par l'IBD1 et l'IBD2. Le test de Wilcoxon pour données appariées donne cependant une p .value < 0.001 ***, ce qui démontre que la différence de notation entre l'IBD1 et l'IBD2 est significative. Il est à noter que ce test, comme tout test non paramétrique, tend à accuser la significativité des différences.

La moyenne des différences de notation obtenues (IBD1-IBD2) est de -0.03, avec un écart-type de 0.33, autrement dit, même si les notes fournies varient significativement d'une version à l'autre, les différences peuvent être considérées comme marginales car elles sont très faibles. La Figure 11 montre d'ailleurs que la médiane des différences est égale à 0 et que le box-plot est très écrasé autour de cette valeur, ce qui témoigne d'une grande similarité de notation entre l'IBD1 et l'IBD2.

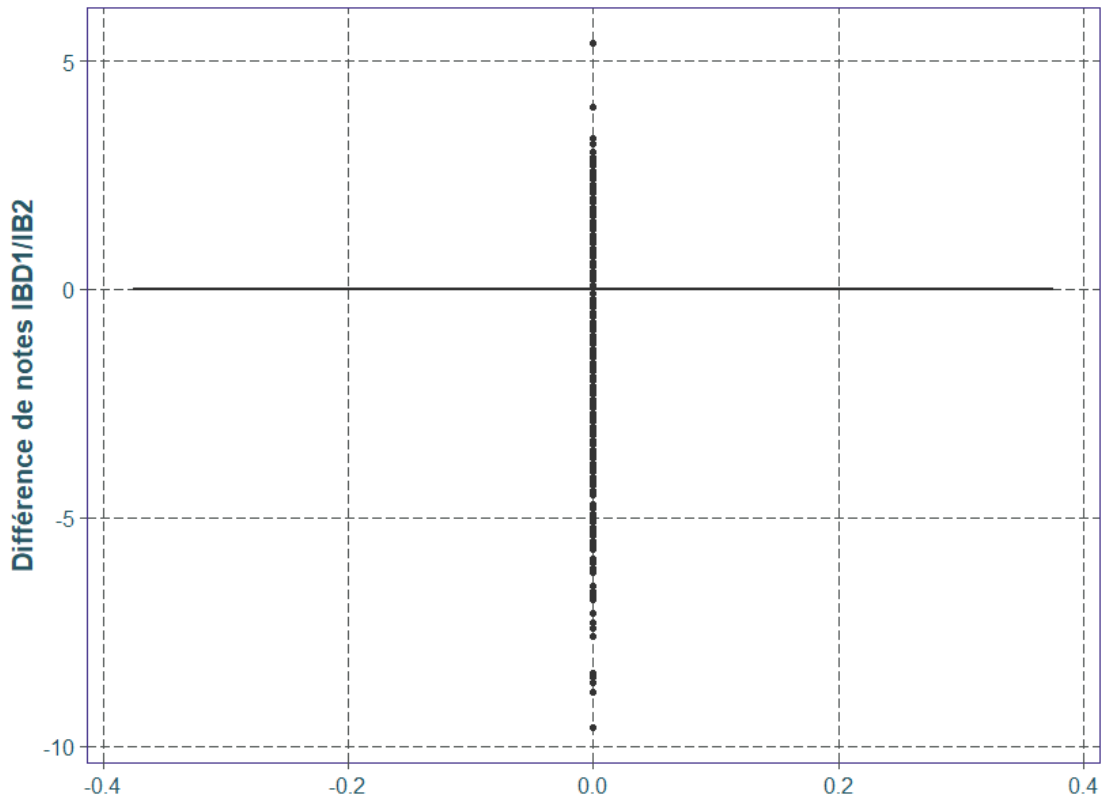


Figure 11: Box-plot de la distribution des différences de résultats entre l'IBD calculé avec l'ancienne (IBD1) et la nouvelle (IBD2) liste de taxons clés

Un test de corrélation de Spearman entre les rangs des notations IBD1 et IBD2 nous a permis d'évaluer une corrélation positive très forte (p .value < 0.001 ***, $R^2 = 0.98$) entre les deux, ce qui démontre une cohérence forte entre les notes fournies par les deux versions de l'IBD. On observe de plus sur la Figure 12 que la relation tend à être linéaire, proche de $y=x$. Il faut noter tout de même que pour certaines valeurs de rang de l'IBD2, les notations obtenues avec l'IBD1 semblent nettement différentes, comme les 'outliers' de la Figure 12 l'indiquent.

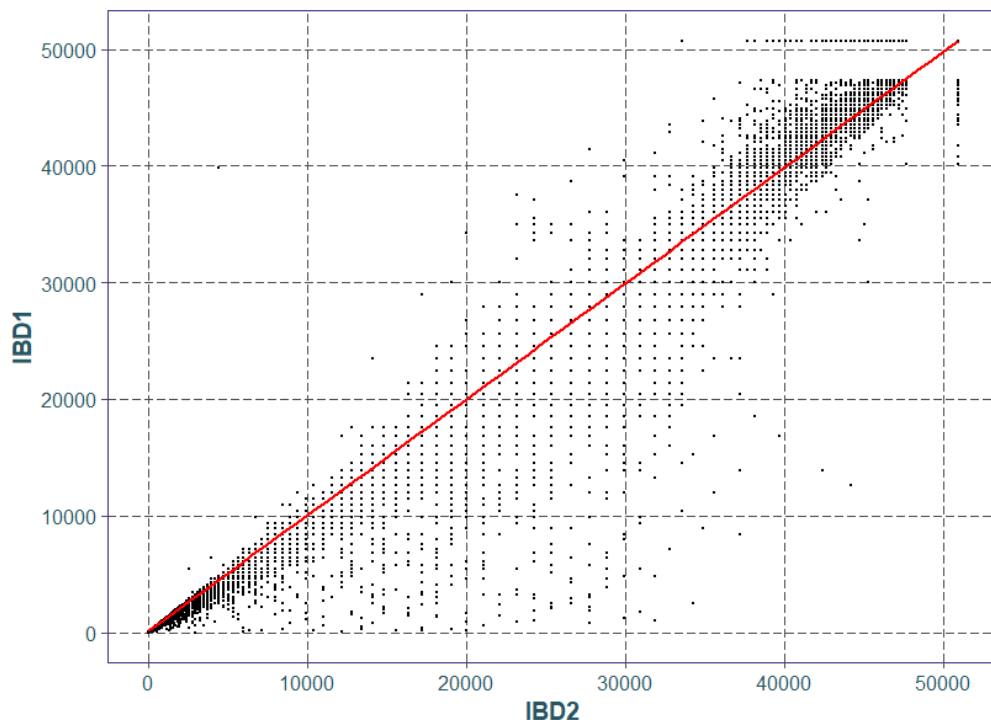


Figure 12: Nuage de points présentant la relation entre les rangs des résultats de l'IBD calculés avec l'ancienne (IBD1) et la nouvelle liste (IBD2) de taxons clés, plus le rang est grand, plus la note IBD est forte (proche de 20).

Pour évaluer cette variabilité de notation de l'IBD2 par rapport à l'IBD1, nous avons estimé la corrélation entre les notes fournies par l'IBD2 et la différence de notation entre IBD1 et IBD2 ($IBD1 - IBD2$). Le test a fourni une p.value < 0.001 *** et un R^2 de 0.3. Il existe donc une corrélation positive faible qui indique que l'IBD2 tend légèrement à dévier les notes des échantillons par rapport à l'IBD1. La Figure 13 témoigne de cette légère variabilité avec 7988 valeurs de différences inférieures à 0 sur 53974 cas au total (15%). 6054 valeurs de différences sont positives (11%), ce qui porte à 74% le nombre de cas où la différence de notation entre l'IBD1 et l'IBD2 est nulle. Les notes sont donc impactées de façon minime par l'ajout des taxons cibles à la liste des taxons contributifs de l'IBD.

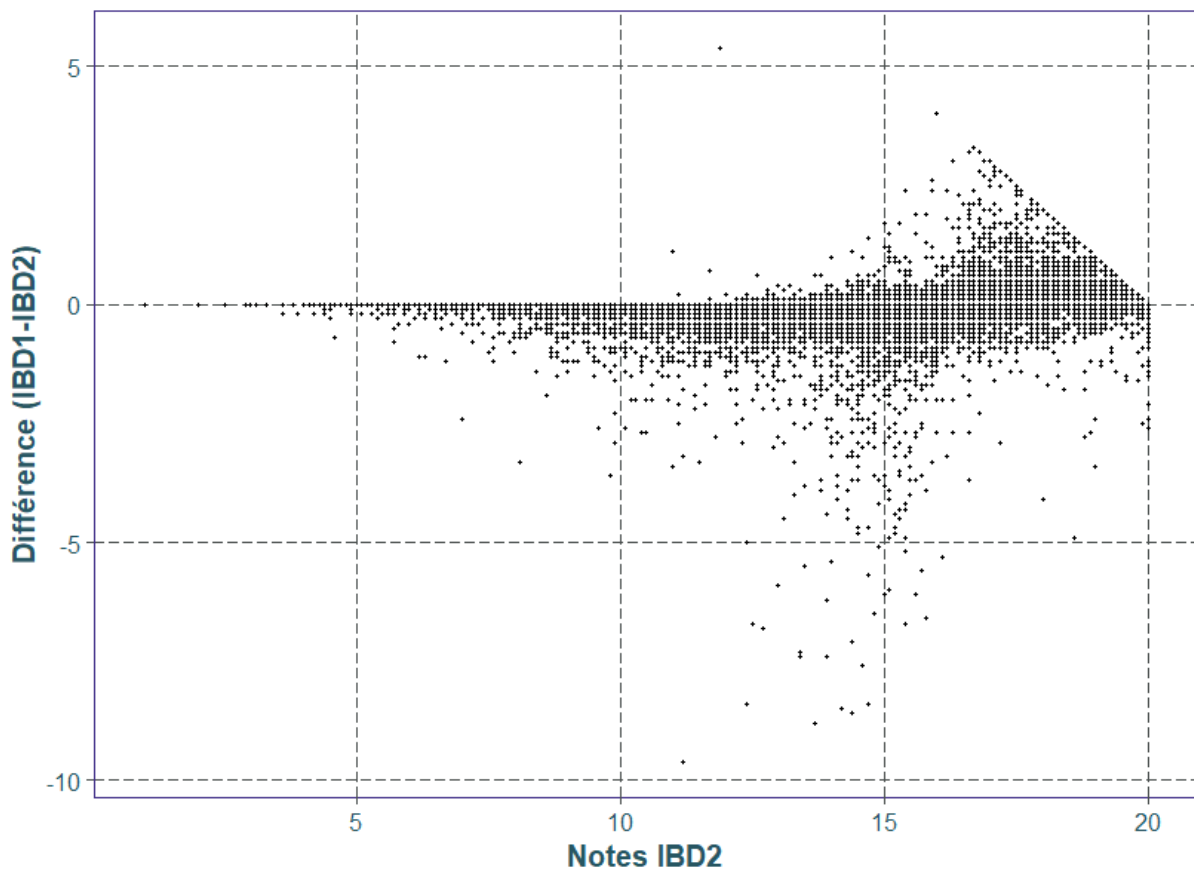


Figure 13: Nuages de points présentant la relation entre les notes fournies par l'IBD2 et la différence de notation entre les deux versions ($IBD1 - IBD2$).

4 CONCLUSION

La méthode d'analyse des taxons indicateurs (IndVal) utilisée dans la présente étude a permis de répondre à la nécessité, exprimée par les experts diatomistes impliqués dans le calcul de l'Indice Biologique Diatomées (IBD), d'associer un profil écologique à des taxons d'intérêt définis au début de ce rapport (voir Tableau 1 et Tableau 2). Au total, 56 taxons ont été ajoutés à la liste des taxons clés de l'IBD. La méthode développée combine analyses statistiques et avis d'expert afin d'obtenir des résultats solides et en phase avec la connaissance de terrain. Cette approche repose sur le fait qu'il n'existe pas de réponse statistique unique permettant d'associer directement le profil écologique d'un taxon à un autre, et que le cas des complexes taxonomiques qui se subdivisent au fil des années en taxons *sensu lato* et *sensu stricto* requiert un avis d'expert maîtrisant l'écologie et la répartition du taxon considéré.

Pour certains taxons, le nombre de relevés disponibles s'est avéré insuffisant pour les inclure de façon fiable dans les analyses. Il faudra attendre que les bases de données soient enrichies au cours des prochaines années pour pouvoir leur attribuer un profil écologique. Des problématiques d'ordre taxonomique nous ont également conduit à écarter certains taxons de la liste finale. Par exemple, le profil écologique que nous avons proposé aux experts diatomistes de l'OFB concernant la nouvelle description de *Cocconeis placentula* (CPLA) (depuis la publication de Jahn *et al* en 2009) était celui de *Cocconeis euglypta* (CEUG). Cependant, nous devons attendre que le SANDRE puisse codifier un CPLA « ancienne version » et un CPLA *sensu* Jahn afin de les distinguer de façon sûre dans les futurs relevés et d'associer le profil de CEUG uniquement au CPLA *sensu* Jahn. De la même façon le profil d'autres taxons présentant la même problématique taxonomique comme *Fragilaria gracilis* (FGRA) et *Gomphonema pumilum* (GPUM) ne sera pas revu, tant que les modifications d'appellation ne seront pas stabilisées et codifiées au niveau du SANDRE.

En conclusion générale de ce travail, la comparaison des notes d'IBD obtenues entre l'ancienne version (basée sur la liste initiale de taxons contributifs) et la nouvelle version a mis en évidence une légère déviation positive de notation pouvant être considérée comme marginale et n'influant pas de manière préoccupante sur la notation des échantillons, et donc sur l'évaluation de l'état écologique des masses d'eau « cours d'eau ».

L'étude que nous présentons dans ce rapport souligne le rôle indispensable de l'expertise et la nécessité de suivre attentivement l'évolution taxonomique des communautés diatomiques pour garantir une gestion fine des écosystèmes aquatiques. Nous pensons à la lumière de ces travaux qu'une révision plus importante des profils écologiques des taxons contributifs de l'IBD est nécessaire dans un futur proche, révision impliquant cette fois un recalcul de l'ensemble des profils. Une nouvelle version de l'IBD sera ainsi générée, et son impact sur l'évaluation de l'état écologique des cours d'eau devra être apprécié.

5 RÉFÉRENCES

- [1] Daily G.C. *Nature's Services: Societal Dependence On Natural Ecosystems*. Island Press, 2012.
- [2] Dudgeon D., Arthington A.H., Gessner M.O., Kawabata Z.-I., Knowler D.J., Lévêque C., Naiman R.J., Prieur-Richard A.-H., Soto D., Stiassny M.L.J., and Sullivan C.A. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, 81, 2006. doi: 10.1017/S1464793105006950.
- [3] Tan X., Dai K., Parajuli K., Hang X., Duan Z., and Hu Y. Effects of phenolic pollution on interspecific competition between microcystis aeruginosa and chlorella pyrenoidosa and their photosynthetic responses. *International Journal of Environmental Research and Public Health*, 16, 2019. doi: 10.3390/ijerph16203947.
- [4] de Carvalho D. R., Flecker A. S., Alves C. B. M., Sparks J. P., and Pompeu P. S. Trophic responses to aquatic pollution of native and exotic livebearer fishes. *Science of the Total Environment*, 681, 2019. doi: <https://doi.org/10.1016/j.scitotenv.2019.05.092>.
- [5] Del Arco A. I., Rico A., and van den Brink P. J. Effects of intra-and interspecific competition on the sensitivity of daphnia magna populations to the fungicide carbendazim. *Ecotoxicology*, 24, 2015. doi: 10.1007/s10646-015-1512-y.
- [6] Buczkó K., Trábert ZS., Stenger-Kovács CS., Tapolczai K., Bíró T., Duleba M., Földi A., Korponai J., Vadkerti E., Végvári ZS., and Ács E. Rapid expansion of an aquatic invasive species (ais) in central- european surface waters; a case study of achnantheidium delmontii. *Ecological Indicators*, 135, 2022. doi: <https://doi.org/10.1016/j.ecolind.2022.108547>.
- [7] Butcher R. W. Studies in the ecology of rivers: Vii. the algae of organically enriched waters. *Journal of Ecology*, 35, 2022. doi: <https://doi.org/10.2307/2256507>.
- [8] Rimet F. and Bouchez A. Life-forms, cell-sizes and ecological guilds of diatoms in european rivers. *Knowledge and Management of Aquatic Ecosystem*, 406, 2012. doi: 10.1051/kmae/2012018.
- [9] Coste M., Boutry S., Tison-Rosebery J., and Delmas F. Improvements of the biological diatom index (bdi): Description and efficiency of the new version (bdi-2006). *Ecological Indicators*, 9, 2009. doi: <https://doi.org/10.1016/j.ecolind.2008.06.003>.
- [10] Debenest T., Pinelli E., Coste M., Silvestre J., Mazzella N., Madigou C., and Delmas F. Sensitivity of freshwater periphytic diatoms to agricultural herbicides. *Aquatic Toxicology*, 93, 2009. doi: DOI: 10.1016/j.aquatox.2009.02.014.
- [11] Almeida S.F.P., Elias C., Ferreira J., Elisabet E., Camilla P., Delmas F., Dörflinger G., Urbanič G., Marcheggiani S., Tison-Rosebery J., Mancini L., and Sabater S. Water quality assessment of rivers using diatom metrics across mediterranean europe: A methods intercalibration exercise. *Science of The Total Environment*, 476-477, 2009. doi: <https://doi.org/10.1016/j.scitotenv.2013.11.144>.

- [12] Wasson J. G., Chandesris A., Pella H., and Blanc L. Les hydro-écorégions : une approche fonctionnelle de la typologie des rivières pour la directive cadre européenne sur l'eau. *Ingénieries eau-agriculture-territoires*, 40, 2004.
- [13] Wetzel C. E., L. Ector, Van de Vijver B., Compere P., and Mann D. G. Morphology, typification and critical analysis of some ecologically important small naviculoid species (bacillariophyta). *Fottea*, 15, 2015. doi: 10.5507/fot.2015.020.
- [14] Jahn R., Kusber W. H., and Romero O. E. *Cocconeis pediculus* ehrenberg and *c. placentula* ehrenberg var. *placentula* (bacillariophyta): Typification and taxonomy. *Fottea*, 9, 2009. doi: 10.5507/fot. 2009.027.
- [15] Krebs C. J. *Ecological methodology*. Benjamin Cummings, 1999.
- [16] Jarman and Angur. *Hierarchical Cluster Analysis: Comparison of Single linkage, Complete linkage, Average linkage and Centroid Linkage Method*, volume 29. 2020. doi: 10.13140/RG.2.2.11388.90240.
- [17] Dufrêne M. and Legendre P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. ecological monographs. *Oecologia*, 67, 1997. doi: [https://doi.org/10.1890/0012-9615\(1997\)067\[0345:SAAI\]2.0.CO;2](https://doi.org/10.1890/0012-9615(1997)067[0345:SAAI]2.0.CO;2).
- [18] Legendre P. and Gallagher E.D. Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129, 2001. doi: <https://doi.org/10.1007/s004420100716>.
- [19] Carayon D., Tison-Rosebery J., and Delmas F. Defining a new autoecological trait matrix for french stream benthic diatoms. *Ecological Indicators*, 103, 2019. doi: <https://doi.org/10.1016/j.ecolind.2019.03.055>.

6 ANNEXES

Annexe 1 : Tableau des occurrences des taxons cibles dans les relevés floristiques

Taxon	2015	2016	2017	2018	2019	2020	2021	Occurrences	Années
ADAM	386	454	330	330	337	338	57	2232	7
ADCV	27	13	21	16	54	73	4	208	7
ADMC	3	38	33	100	99	513	220	1006	7
ADMI	3829	3785	3531	3522	3753	3608	528	22556	7
ADMO	406	624	587	703	859	882	269	4330	7
ADRU	318	296	412	321	337	367	52	2103	7
ADSK	absent	absent	absent	1	absent	6	absent	7	2
ADTC	absent	absent	absent	absent	absent	32	absent	32	1
AFBA	absent	absent	absent	absent	17	13	absent	30	2
AHOF	18	64	64	171	244	324	33	918	7
ALBL	240	304	283	282	312	253	16	1690	7
AMCD	6	13	86	66	205	285	37	698	7
AZHA	32	40	29	26	52	72	19	270	7
CBPY	absent	absent	absent	absent	4	absent	absent	4	1
CEXF	57	58	50	45	60	50	43	363	7
CLTL	12	14	7	1	5	6	5	50	7
CPLA	983	1045	1061	937	1160	1252	136	6574	7
CROX	absent	absent	absent	absent	3	3	absent	6	2
CSBH	3	5	3	3	9	3	4	30	7
CSUT	7	4	2	5	4	9	9	40	7
CTDE	absent	1	absent	5	5	11	4	26	5
EARB	17	13	11	12	8	18	9	88	7
EBNA	absent	absent	3	9	11	16	absent	39	4
ECAL	absent	absent	absent	4	12	18	17	51	4
ECTO	absent	absent	absent	absent	1	1	absent	2	2
EJUE	absent	absent	absent	2	1	4	absent	7	3
ENEE	1	4	2	1	2	absent	2	12	6
ENSI	9	30	16	48	78	135	7	323	7
EPSG	1	1	3	2	4	4	absent	15	6
FCAD	absent	2	10	65	141	135	24	377	6
FGRA	602	737	667	600	704	628	129	4067	7
FMIV	1	3	17	68	127	120	14	350	7
FNEV	33	52	49	40	50	48	8	280	7
FNIN	9	14	26	19	40	37	9	154	7
FPDE	absent	absent	8	10	11	6	10	45	5
FPRU	188	206	266	239	296	259	54	1508	7
GAGV	83	106	85	70	60	52	45	501	7
GAUR	8	10	13	21	22	20	2	96	7
GCUN	189	231	230	198	256	235	31	1370	7
GINN	3	7	16	31	28	25	absent	110	6
GMIS	5	8	15	16	18	22	9	93	7
GPPY	absent	absent	absent	absent	absent	absent	absent	0	0

GPUM	433	358	354	421	470	599	71	2706	7
GTNO	9	25	14	16	15	14	15	108	7
GVRD	24	40	46	42	41	46	absent	239	6
HPDA	67	88	102	66	98	109	13	543	7
HRUT	8	3	3	absent	1	4	absent	19	5
HTHU	12	10	6	7	5	5	7	52	7
LHLU	absent	3	5	17	23	22	6	76	6
NDDF	165	127	157	166	145	172	10	942	7
NFSO	3	2	52	128	116	141	10	452	7
NPHP	3	2	5	5	9	5	1	30	7
NSBN	24	16	15	10	8	10	8	91	7
NSTS	901	945	901	890	825	923	108	5493	7
NYCO	502	561	666	426	542	469	74	3240	7
PLRC	5	6	8	5	5	13	6	48	7
PMNT	95	68	108	92	119	111	2	595	7
POVA	absent	3	19	46	71	77	10	226	6
PSPO	absent	1	absent	6	6	2	3	18	5
PSXO	5	5	19	33	49	58	3	172	7
SBOS	absent	absent	absent	1	2	3	1	7	4
SCAN	11	34	28	45	45	48	6	217	7
SCRA	5	3	14	38	46	75	absent	181	6
SEAT	32	2	4	661	1262	1557	122	3640	7
SECA	absent	1	3	6	7	16	1	34	6
SESP	absent	absent	absent	10	47	41	98	3	
SNIG	3193	3208	3271	3144	3135	3056	371	19378	7
SPDV	52	77	118	151	149	188	3	738	7
SRAE	17	14	10	31	43	42	absent	157	6
SRBU	4	12	10	8	5	6	1	46	7
SSBG	34	28	26	40	63	56	3	250	7
