HAL
open science

Incertitudes dans les mesures macrophytes en cours d'eau : Impact des confusions de détermination taxinomique dans l'évaluation de l'état écologique

M. Jeanty

To cite this version:

M. Jeanty. Incertitudes dans les mesures macrophytes en cours d'eau : Impact des confusions de détermination taxinomique dans l'évaluation de l'état écologique. Sciences de l'environnement. 2015. hal-02606083

HAL Id: hal-02606083

https://hal.inrae.fr/hal-02606083v1

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.

Ecole des Métiers de l'Environnement
Campus de Ker Lann – Rennes
35170 BRUZ
Tél. : 02.99.05.88.00
www.ecole-eme.fr

Formation Ingénieur GIE en 5 ans

Auteur : JEANTY Manon

I5 - 5ème année cycle ingénieur
Année : 2014-2015
Correcteur : DJELAL Hayet
Version N°1 du.11/09/2015

STAGE DE FIN D'ETUDE

Incertitudes dans les mesures « macrophytes en cours d'eau »
-
Impact des confusions de détermination taxinomique
dans l'évaluation de l'état écologique

CONFIDENTIEL

Centre de Bordeaux
Tel: +33 (0)5.57.89.08.00
50 avenue de Verdun - Gazinet
33612 Cestas Cedex
http://www.irstea.fr/accueil

Résumé

Mots-clefs : confusion, DCE, IBMR, incertitudes, macrophytes

Pour répondre à l'exigence de la Directive Cadre sur l'Eau (DCE) prescrivant de maitriser la « confiance » et la « précision » lors de l'évaluation de la qualité des cours d'eau, l'étude des incertitudes des méthodes de surveillance et d'évaluation est nécessaire. Actuellement peu d'études existent, en particulier en ce qui concerne l'incertitude liée à la confusion taxinomique. Le présent travail s'est intéressé à l'impact des confusions de détermination taxinomique sur la méthode française d'évaluation à partir des macrophytes aquatiques, basée sur l'IBMR (indice biologique macrophytique en rivière). Nous nous sommes basés sur une démarche de consensus entre experts pour créer une matrice de confusion. Cette matrice, élaborée à une échelle nationale, a été couplée à une liste floristique issue des réseaux de surveillance DCE afin de générer de façon aléatoire, et selon les coefficients de confusion attribués par les experts, des listes floristiques comprenant des confusions taxinomiques. Ce protocole d'analyse permet de mesurer l'impact que peuvent avoir ces confusions taxinomiques, qui ont lieu dans la réalité, sur l'évaluation de la qualité des eaux. Les IBMR et classes d'état écologique obtenus ont alors pu être comparés aux IBMR et classes d'état originaux. Nous avons obtenu tout un panel d'écarts à l'origine et ainsi pu déterminer les couples impactant le plus l'IBMR et les classes d'état écologique, ainsi que les différences liées aux biotypes et régions où se situent les sites de surveillance. Ce travail a pour objectif de contribuer à identifier les sources d'incertitude des méthodes d'évaluation et estimer leur impact, dans le but de proposer des protocoles de prise en compte de ces niveaux de confiance ou des ajustements correctifs dans les méthodes nationales prescrites par les règlements pour l'évaluation de l'état écologique des cours d'eau.

Abstract

Keywords : confusion, DCE, IBMR, uncertainties, macrophytes

The Water Framework Directive (WFD) request reliability and precision in water stream quality monitoring programs. Meeting both requirements imply the study of uncertainties derived from monitoring and assessment methods. Currently, little scientific studies exist, especially about the uncertainties due to taxonomic confusion. This work is about the impact of the taxonomic identification confusions on the French method for assessment from aquatic macrophytes, based on the IBMR (Macrophyte biological index for rivers). We have used a consensus approach among experts to generate a confusion matrix. This matrix, made on a national scale, has been coupled to a floristic list from the WFD monitoring network. Hence, floristic lists containing taxonomic confusions can be generate randomly and accordingly to the confusion coefficients assigned by the experts. This analysis protocol allows to measure the impact these taxonomic confusions have on the water quality assessment. The resulting IBMR and ecological status classes were compared with the original ones. We obtained a wide range of differences with the initial reports. Subsequently, on the one hand we were able to figure out the most impacting couples for the IBMR and the ecological status classes, on the other hand the variances related to biotypes and areas where the monitoring sites are located. This study aims to contribute to the identification of the uncertainties sources from assessment methods, and the assessment of their impact. The ultimate goal of the study is to propose protocols that take into account these reliability levels, or corrective adjustments in the national methods required by the regulations for the assessment of the ecological states of the streams.

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage et qui m'ont aidée lors de la rédaction de ce rapport.

Tout d'abord, j'adresse mes remerciements à mon professeur, Mme Hayet GOUILLAUD de l'Ecole des Métiers de l'Environnement qui s'est occupée du suivi de mon stage et s'est assurée du bon déroulement de celui-ci.

Je tiens à remercier vivement mon maitre de stage, M. Christian CHAUVIN, ingénieur de recherche spécialisé en phytoécologie aquatique à l'Irstea de Bordeaux, pour son accueil, le temps passé ensemble, et le partage de son expertise au quotidien.

Je remercie également toute mon équipe avec Nina DAGENS, Ingénieur d'étude Aquaref, Thibaut FERET, Ingénieur d'études - Biostatistiques, bioindication, hydrobiologie, et Sandrine LORIOT, ingénieur d'étude spécialisée en phytoécologie aquatique, pour leur accueil, leur esprit d'équipe, et leur aide précieuse pour m'aider à avancer dans ma mission, mais aussi à la réalisation du mémoire.

Je tiens tout particulièrement à remercier tous les experts ayant participé à cette mission, sans lesquels tout ce travail n'aurait pas pu se faire :

- Monique Aubert, Hydrobiologiste consultante indépendante
- Marie Christine PELTRE, Université de Lorraine-Metz
- Christophe Laplace-Treyture , Irstea Bordeaux
- Julie Coudreuse, Agrocampus Ouest de Rennes
- Elisabeth Lambert, Université Catholique de l'Ouest - Angers
- Jacques Haury, INRA- Agrocampus Ouest de Rennes

Je remercie de même toutes les personnes du centre Irstea Bordeaux, auprès de qui j'ai pu passer 6 mois très enrichissants autant professionnellement que du point de vue personnel.

SOMMAIRE

Introduction 1

I- Analyse de la situation initiale 4

1) Contextualisation 4

a. Cadre réglementaire 4

b. Les bioindicateurs, des outils d'évaluation de l'état écologique 5

c. Les macrophytes en tant que bioindicateurs 5

2) L'incertitude en hydrobiologie 10

a. Programme « incertitudes » Irstea – ONEMA 10

b. L'incertitude dans la littérature scientifique 12

c. Les travaux de J. Wiederkehr 14

3) Résultats obtenus 15

II- Outils et démarche scientifique 16

1) La base de données PANDORE 16

2) Fichier répertoire des couples sujets à confusion 16

3) Matrice de confusion 18

a. Principe et contenu de la matrice de confusion 18

b. Constitution du fichier de consultation des experts 18

c. Consultation des experts 20

d. Homogénéisation des réponses des experts 21

e. Synthèse des avis des experts et création de la matrice 22

4) La dégradation des listes 24

5) Aides à la décision dans la démarche 25

III- Résultats et perspectives 30

1) Répertoire de l'analyse bibliographique 30

2) Description de la matrice de confusion taxinomique 30

3) Impact des confusions par approche unitaire 32

4) Impact des confusions par approche multiple 34

a. Etude d'impact global sur l'IBMR 34

b. Impact en termes d'évaluation de l'état écologique 36

c. Régionalisation de l'impact sur le classement 37

5) Evaluation de l'étude 38

a. Discussion 38

b. Bénéfices environnementaux, économiques et sociétaux des opérations 39

c. Analyse des écarts des résultats 39

d. Devenir des résultats 39

Conclusion 40

Glossaire 42

Références bibliographiques 43

Fiche compétences 45

Annexes 47

Introduction

Dans le cadre de ma formation d'ingénieur en Génie Industriel de l'Environnement à l'Ecole des Métiers de l'Environnement (EME), j'ai pu effectuer mon Projet de Fin d'Etudes au sein d'un centre de Recherche : l'Institut National de Recherche en Science et Technologies pour l'Environnement et l'Agriculture (IRSTEA). Ce stage s'est déroulé sur une durée de 6 mois, du 16 mars au 11 septembre 2015 au Centre de Bordeaux (Gazinet-Cestas - 33).

L'IRSTEA est un organisme de recherche représenté par 9 centres en France, rassemblant 1604 collaborateurs dont 1219 chercheurs, ingénieurs, doctorants et post-doctorants. Toutes ces personnes se répartissent dans 3 départements de recherche : eaux, écotechnologies, territoires.

Face aux enjeux environnementaux, les décideurs politiques, acteurs économiques et sociétaux nécessitent des connaissances scientifiques apportant résultats et/ou solutions opérationnelles, que ce soit sous forme d'expertise ou d'innovation. C'est là que l'Irstea entre en jeu, en tant qu'organisme national de recherche, il va apporter expertise et appui aux politiques publiques.

L'Irstea travaille actuellement à l'échelle du territoire, et vise à étendre son modèle de recherche finalisée en le mettant également au service des politiques européennes.

Ce stage s'insère dans les missions de l'unité "Ecosystèmes aquatiques et changements globaux" (EABX) [1] dont les recherches traitent de l'évaluation des réponses des écosystèmes aquatiques continentaux (estuaires, lacs, rivières) ou espèces aquatiques aux pressions anthropiques (pêche, fragmentation des milieux, contamination, changement climatique) comme des mesures de restauration.
L'unité EABX a pour objectif l'identification des domaines de viabilité de ces différents systèmes pour l'aide à la prise de décision publique afin de rétablir un bon état de ces écosystèmes et de participer à la conservation des espèces menacées. [1]
L'ingénierie qui en ressort permet [1] :

- L'évaluation de l'état écologique des masses d'eaux :
  - estuariennes par l'étude des communautés de poissons
  - rivières et lacs par l'étude des communautés de végétaux aquatiques
- La restauration des espèces de poissons migrateurs amphihalins (anguille, alose, esturgeon, saumon en particulier) par le développement des méthodes en matière de biologie de la conservation.

Au sein de l'unité EABX se trouve l'équipe CARMA [2], spécialisée dans le fonctionnement et les réponses des écosytèmes aquatiques (macrophytes, phytobenthos, et phytoplancton) sous l'influence des altérations anthropiques des cours d'eau et plans d'eau.
C'est au sein de cette équipe que mon stage s'est déroulé.

A partir de 1970, la qualité de l'eau est devenue un enjeu prioritaire dans la politique de l'Union Européenne. L'intérêt premier était porté sur les usages anthropiques de l'eau (eau potable, baignade, pisciculture, conchyliculture), puis il s'est de même tourné vers la réduction des pollutions (eaux usées, nitrates d'origine agricole). [3]

L'approche de la problématique de la gestion de l'eau à l'échelle européenne est logique de par le fait que les cours d'eau ne se limitent pas aux frontières et peuvent parcourir plusieurs pays (comme le Rhin, la Meuse, la Sambre, l'Escaut et le Rhône). De même une protection des eaux de mer s'applique à travers des conventions internationales. [3]

De ce fait, une trentaine de directives sur l'eau ont ainsi été conçues au sein de la législation européenne [3]. Parmi celles-ci se trouve la Directive Cadre sur l'Eau (DCE) qui s'inscrit au niveau européen, et qui fixe aux États membres l'objectif d'atteindre le bon état écologique des masses d'eau.

En effet, l'Union européenne a élaboré un « cadre commun pour la protection et la gestion de l'eau, ayant pour objectifs de protéger et de rétablir la qualité écologique des écosystèmes aquatiques, et de garantir à long terme une utilisation durable de l'eau aux personnes, aux entreprises et à la nature ». [4]
La DCE vise ainsi plusieurs finalités comme « la prévention et la réduction de la pollution, la promotion d'une utilisation durable de l'eau, la protection de l'environnement, l'amélioration de l'état des écosystèmes aquatiques et l'atténuation des effets des inondations et des sécheresses ». [4]

Pour ce faire, des outils sont utilisés afin d'estimer la qualité des eaux. Il s'agit des indicateurs, et plus précisément ici des bioindicateurs (indicateurs biologiques), qui vont permettre de mesurer la qualité des différentes masses d'eau. Au cours de ce stage, c'est l'Indice Biologique Macrophytique en Rivière (IBMR), l'indicateur national français basé sur les végétaux aquatiques en cours d'eau, qui a été utilisé. Cet indice contribue à la détermination de l'état écologique des cours d'eau en répondant aux prescriptions de la DCE.

Cependant la détermination de cette qualité est soumise à tout un panel d'incertitudes apparaissant à chaque étape de la démarche de l'évaluation de l'état écologique.
De ce fait, la DCE demande la détermination d'une « confiance » et d'une « précision » lors de cette évaluation de la qualité des cours d'eau. Il s'agit donc d'un impératif réglementaire.

En définitive, ce qui est important au sens de la DCE, dans l'étude de l'incertitude, est la bonne classification de l'état écologique des eaux. Une estimation précise de cette incertitude, pour chaque type de cours d'eau, fournira aux gestionnaires une aide à la décision pour que des actions pertinentes soient mises en place, et que les moyens économiques à investir soient adaptés aux besoins réels.

Un groupe de travail piloté par Irstea a permis de rassembler les différents partenaires d'Aquaref [5] (Onema, Ineris, Irstea, Ifremer) autour de la problématique des incertitudes dans les mesures hydrobiologiques [6]. Les recherches ont dans un premier temps été orientées vers les incertitudes liées à l'acquisition de la donnée, soit les incertitudes sur la mesure. A l'instar du programme mis en place dès 2009 sur les invertébrés aquatiques (Archambault 2012), un programme de R&D a été mis en œuvre en partenariat avec l'ONEMA, visant à identifier et estimer les incertitudes liées à l'acquisition des données de surveillance « macrophytes cours d'eau » et « diatomées cours d'eau ».
Au cours de ce stage, nous nous sommes donc penchés sur l'incertitude liée à la confusion taxinomique, et son impact sur l'IBMR et l'évaluation de l'état écologique des cours d'eau.

Actuellement, peu de travaux ont été faits sur les incertitudes en hydrobiologie, et un seul a traité de l'impact des confusions liées à l'identification taxinomique sur les valeurs d'IBMR. Notre travail a donc permis de comparer les résultats de l'approche de ces travaux antérieurs récents à nos propres conclusions.

Le but de cette mission est donc de pouvoir répondre aux questions suivantes :
Les erreurs d'identification effectuées lors de relevés macrophytiques en cours d'eau entraînent-elles des variations importantes des indicateurs d'état écologique en jeu? Quels sont les aspects de l'exploitation des données les plus impactés par ces confusions ? Quelles pistes pour gérer cette source d'incertitude ?

Ce rapport se présente sous 3 parties.
Tout d'abord le contexte est exposé afin de bien définir la réglementation en place, les métriques et indicateurs utilisés, mais aussi la problématique que présentent les incertitudes, et les travaux associés qui ont été effectués à ce jour
Par la suite ont été développés le matériel et les méthodes utilisés, dans chaque étape de la démarche, pour l'étude de l'impact de la confusion taxinomique.
Finalement, chacun des résultats obtenus a été analysé dans le but de répondre à la problématique, et de discuter de l'intérêt et des perspectives à envisager en lien avec cette démarche et les résultats obtenus.

I- Analyse de la situation initiale

1) Contextualisation

a. Cadre réglementaire

La Directive Cadre sur l'Eau (DCE) (WFD, 2000/60/EC) [7] a été adoptée le 23 octobre 2000 par le Parlement Européen et le Conseil. Centrée sur la préservation du milieu naturel et associée à une obligation de résultats, elle implique la planification et la mise en œuvre d'actions de diagnostic et de restauration à différents niveaux (état, bassin, SAGE,...).
Elle établit, pour une politique communautaire, un cadre pour la gestion et la protection des eaux par grand bassin hydrographique au plan européen.
La DCE joue un rôle stratégique et fondateur en ce qui concerne la politique de l'eau.

Les grands objectifs de la DCE sur le plan écologique se résument sous les points principaux suivants :

- Obtenir pour 2015 des eaux de classe de qualité de "bon état" (selon les aspects écologique et chimique)
- Préserver les milieux aquatiques définis comme en "très bon état".
- Obtenir pour les milieux fortement artificialisés un bon potentiel d'état.
- Supprimer avant 2020 les rejets des 41 substances dangereuses prioritaires.

L'état écologique, au sens de la DCE, est défini comme « *l'expression de la qualité de la structure et du fonctionnement des écosystèmes aquatiques associés aux eaux de surface ».[8]*

Le bon état écologique est défini, quant à lui, lorsque « les valeurs des éléments de qualité biologique […] montrent de faibles niveaux de distorsion résultant de l'activité humaine, mais ne s'écartent que légèrement de celles normalement associées à ce type de masse d'eau de surface dans des conditions non perturbées » (DCE, annexe V : Définitions normatives des classifications de l'état écologique). [8]

Cette limite « inférieur au bon état » comporte un enjeu fort, car, pour toutes les masses d'eau qualifiées inférieures au bon état écologique ou chimique, l'Etat-membre est dans l'obligation de mettre en place un programme de mesures correctives, et de montrer le retour au bon état.
Les mesures prises sont majoritairement de nature technique ou réglementaire, mais peuvent aussi constituer des incitations financières ou encore des accords volontaires.

La DCE introduit également l'obligation de mettre en place différents réseaux de surveillance [9] de l'état des masses d'eau. Cette surveillance est encadrée précisément par des règlements nationaux, et comporte, entre autres, des prescriptions méthodologiques (paramètres à suivre, protocoles, fréquence de mesure, etc.) et de maîtrise de la qualité des données produites (agrément des laboratoires, cf EN17025 [10]).

b. Les bioindicateurs, des outils d'évaluation de l'état écologique

Les organismes vivants intègrent les contraintes environnementales de leur habitat. [11]
De ce fait, la présence d'une espèce, mais aussi son recouvrement et sa biomasse, vont dépendre de sa possibilité de dispersion, de colonisation, et de son adaptation aux conditions de son environnement qui pourront la faire proliférer, ou régresser voire disparaître.[12]

De cette manière, l'apparition d'une substance polluante, ou la perturbation de l'hydromorphologie, dans un milieu va entraîner des modifications dans les peuplements végétaux.

Une espèce, ou un groupe d'espèce, est considérée comme bioindicatrice lorsque ses caractéristiques (biologiques, morphologiques) témoignent de l'état d'un système écologique, et ainsi fournissent une indication sur le niveau de dégradation éventuelle du milieu. [11] [13]

Grâce à l'utilisation de bioindicateurs, il devient possible d'analyser la nature et l'intensité des contraintes subies par les différents taxons présents. Les indices hydrobiologiques sont donc un outil bien adapté à l'évaluation globale de la qualité du milieu, et donc de l'eau, intégrant la complexité de l'écosystème, mais il sera souvent difficile de relier précisément ces informations aux métriques de l'environnement prises séparément. [11]

c. Les macrophytes en tant que bioindicateurs

Parmi les compartiments biologiques à surveiller selon les exigences de la DCE, on trouve les macrophytes.
Les macrophytes aquatiques regroupent tous les végétaux aquatiques ou amphibies (ou leur colonie) visibles à l'œil nu. Cette définition se veut très fonctionnelle et opérationnelle et ne correspond en aucun cas à une classification botanique en tant que telle. Elle prend ainsi en compte des phanérogames (hydrophytes et hélophytes), des ptéridophytes, des bryophytes (hépatiques et mousses), des lichens, des macro-algues (characées), ainsi que des colonies macroscopiques en thalles ou paquets de filaments d'algues, de cyanobactéries fixées, de bactéries et champignons hétérotrophes. (cf figure 1 et tableau 1) [11] [14]

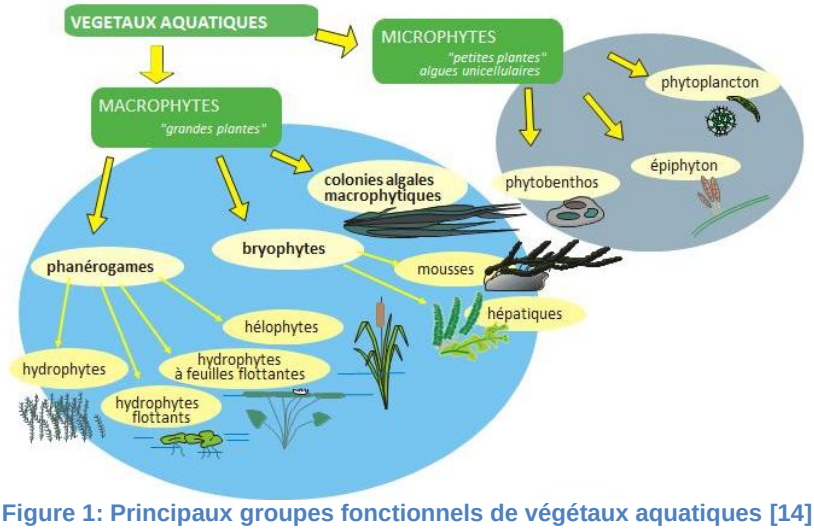
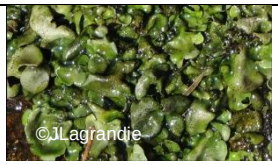


Figure 1: Principaux groupes fonctionnels de végétaux aquatiques [14]

Les macrophytes sont des organismes sensibles à la minéralisation, au niveau trophique global de leur milieu, ainsi qu'aux conditions hydromorphologiques de celui-ci.
De ce fait, les groupements végétaux seront différents d'un écosystème lentique (eaux douces à circulation lente ou nulle) à un écosystème lotique (eaux courantes).

Tableau 1 : Exemples d'espèces/genres appartenant aux groupes taxinomiques composant les macrophytes

Groupe taxinomique	Photo d'une espèce appartenant au taxon	Nom de l'espèce/genre
Algue		*Melosira sp.*
Bryophyte		*Fontinalis antipyretica*
Characée		*Nitella mucronata*
Hétérotrophe		*Sphaerotilus sp.*
Lichen		*Dermatocarpon weberi*
Phanérogame		*Potamogeton nodosus*
Ptéridophyte		*Equisetum ramosissimum*

L'indice biologique

Au début des années 2000, J. Haury (Agrocampus de Rennes) et le GIS Macrophytes (cf *II.3.c. « Consultation des experts »*) ont développé un Indice Biologique à partir des communautés Macrophytiques en Rivières [15]. Cet indice, l'IBMR, a été construit dans l'objectif de répondre au niveau trophique global du milieu. Il s'agit d'un indice biocénotique qui se base sur un inventaire complet de la flore aquatique observée, à partir duquel est calculé un indicateur traduisant une affinité moyenne du peuplement pour un niveau trophique donné.

Cet indice est normalisé en France depuis 2003 [11], et constitue le protocole officiel en vigueur pour l'échantillonnage des macrophytes dans le cadre des réseaux de surveillance des cours d'eau mis en place par la DCE [14]

Cet indicateur est à la base de la méthode d'évaluation développée par l'Irstea pour les besoins de l'évaluation DCE.
Le protocole de relevé est compatible avec la norme européenne NF EN 14184– Guide pour l'étude des macrophytes aquatiques dans les cours d'eau. [16]
Cette norme est actuellement composée des éléments suivants [11, 14] :
– un protocole de relevé sur site des macrophytes présents sur la station de mesure,
– une liste des taxons pris en compte pour le calcul de l'IBMR et les valeurs caractéristiques (côte spécifique CS, coefficient de sténoécie E)
– une annexe pour l'adaptation du protocole aux grands cours d'eau non prospectables à pied.

Le calcul de l'IBMR [11, 14] correspond à une moyenne des cotes spécifiques pondérées par l'abondance et le coefficient de sténoécie de chaque taxon :

$$IBMR = \frac{\sum_{i}^{n} E_i \times K_i \times CS_i}{\sum_{i}^{n} E_i \times K_i}$$

Avec :

- CS (0-20) : la cote spécifique traduisant l'affinité pour les conditions trophiques du milieu, allant de 0 (dystrophe/hypereutrophe et pollution organique forte) à 20 (très oligotrophe).
- E (1-3) : Le coefficient de sténoécie reflétant l'amplitude écologique et donc la représentativité par rapport à des conditions mésologiques précises. Une valeur de 1 décrit un taxon euryèce (espèce à grande valence écologique, c'est-à-dire pouvant coloniser de nombreux habitats, car supportant d'importantes variations de facteurs écologiques), et une valeur de 3 un taxon sténoèce (espèce à faible valence écologique).
- K (0-5): la classe d'abondance représentant le recouvrement spécifique, classé de 1 (abondance la plus faible) à 5 (abondance la plus élevée).
- n : le nombre de taxons dans la liste floristique

Les cotes et coefficients spécifiques sont précisés dans la liste de référence, comportant 208 taxons, fournie par la norme. [11]

Le score IBMR varie entre 0 (dystrophe, très forte charge en matières organique) et 20 (état ultra-oligotrophe : le milieu est particulièrement pauvre en éléments nutritifs).

Les classes de niveau trophique proposées dans la norme ne constituent pas des classes de qualité utilisables dans le cadre de l'évaluation DCE, mais elles peuvent tout de même servir à titre d'indication ou de diagnostic.
Un guide d'application en cours d'élaboration par l'AFNOR [11] précise les moyens de mise en œuvre du protocole standard, et décrit les protocoles à suivre pour l'échantillonnage des grands cours d'eau.

L'homogénéité et la pertinence au niveau national du protocole suivi étant importantes, la mise en place de formations professionnelles a été nécessaire, et permet ainsi de diminuer la variabilité inter-opérateurs pouvant exister lors de l'application de la méthode. [17]

Les conditions du milieu étant constamment en changement, l'IBMR va forcément être plus ou moins influencé par les différentes contraintes biotiques et abiotiques du milieu étudié.
Comme pour n'importe quel indice basé sur un peuplement en milieu naturel ou assimilé, il parait évident que les différentes caractéristiques biotiques et abiotiques du milieu vont interagir ensemble de façon complexe, et il ne sera pas possible de connaître l'influence exacte de chacune de ces caractéristiques, et les réponses univoques de l'IBMR.
Cependant, il est possible de donner une idée de manière qualitative de la réponse attendue de l'IBMR (cf tableau 2).
Les indications suivantes correspondent à des tendances globales, les situations locales ayant des conditions environnementales différentes, et donc des facteurs biotiques/abiotiques prépondérants qui diffèrent par rapport à d'autres situations. De plus, ces facteurs sont corrélés entre eux, à différents degrés. [11]

Tableau 2 : influence relative des principaux facteurs mésologiques sur la réponse de l'IBMR. [11]

métriques	influence
Charge en nutriments	+ + + +
Eclairement	+ +
Température	++
Pollution organique	+ + + +
Hydrologie	+
Type de substrat	+ +
Mobilité du substrat	+ + +
Diversité des habitats	+

La méthode française d'évaluation de l'état écologique

Le système d'évaluation « macrophytes en cours d'eau » a été développé par Irstea en 2011. Il comprend les valeurs de référence de l'IBMR attribuées aux différents types nationaux de cours d'eau (cf Annexe 1), permettant ainsi le calcul de l'EQR (Ecological Quality Ratio), ainsi que les bornes de classes d'état à l'aide de ce ratio (cf figures 2 et 3 et tableau 3). [18]
Ces IBMR de référence sont issus des sites de référence, qui sont des sites sélectionnés pour chaque type de cours d'eau, dans le but d'établir des valeurs de référence du "bon état". Celles-ci sont alors utilisées pour le calcul des différents indicateurs de qualité biologique (invertébrés, diatomées, poissons, macrophytes). [19]

L'EQR se calcule de la façon suivante : $$EQR = \frac{IBMR}{IBMR\ ref}$$

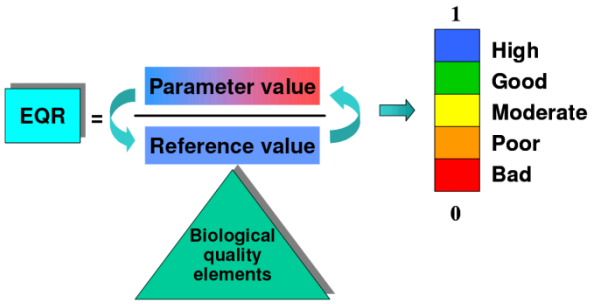


Figure 2: Les classes de qualité définies selon l'EQR [20]

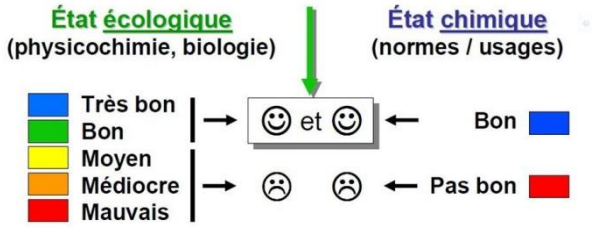


Figure 3 : Les classes d'état écologique et chimique selon les demandes de la DCE [13]

Les limites de classes d'état, définies à partir l'EQR dans un contexte d'utilisation de l'IBMR, se définissent selon les limites de classes officielles présentées ci-dessous.

Tableau 3: Valeurs inférieures des limites des classes d'état, exprimées en EQR, par type de cours d'eau pour l'IBMR [21]

ÉLÉMENTS DE QUALITÉ	INDICE	LIMITES DES CLASSES D'ÉTAT IBMR en EQR			
		Très bon / Bon	Bon / Moyen	Moyen / Médiocre	Médiocre / Mauvais
Macrophytes	IBMR	0,92	0,77	0,64	0,51
Les valeurs de l'IBMR figurant dans ce tableau ont pris en compte la décision de la commission du 20 septembre 2013 relative à l'inter-étalonnage.					

La limite de classe entre Bon et Moyen est la limite la plus importante au sens de la DCE, bien que la moins certaine du point de vue écologique. En effet la classe de Bon état est la classe minimum visée pour tous les cours d'eau afin de répondre aux demandes de la DCE. De ce fait, les Etats doivent atteindre cet état de qualité, en investissant dans des actions afin de viser cet état. Ainsi il est essentiel d'avoir une précision suffisante en particulier au niveau de cette limite.

Afin de répondre aux exigences de la DCE, une certaine fiabilité est attendue quant aux classes obtenues pour chaque cours d'eau. Il en ressort donc une nécessité d'élargir les connaissances et compétences dans le domaine des incertitudes en hydrobiologie.

2) L'incertitude en hydrobiologie

Certaines variabilités vont créer une incertitude dans la prise de ces mesures ainsi que dans le calcul de ces différentes métriques. Une étude bibliographique a permis de mieux appréhender cette problématique, et de savoir où en était l'avancée des connaissances et méthodes à ce sujet.

a. Programme « incertitudes » Irstea – ONEMA

Le terme d'incertitude n'est pas employé tel quel dans la DCE (Directive 2000/60/CE) [22], il s'agit plutôt des concepts de « confiance » et « de précision » qui y sont exprimés (annexes 2 et 5 de la Directive).
Ce qui est réellement demandé par la DCE:

- « pour les valeurs de référence, un niveau de confiance suffisant du réseau et des méthodes »,
- « une confiance et une précision suffisantes dans la classification des éléments de qualité »,
- « une estimation du niveau de confiance et de précision des programmes de surveillance et des résultats obtenus ».

Cependant, il n'est à aucun moment question d'un niveau exigé par la DCE ou encore de méthodologie d'évaluation de cette confiance et précision[6]. Chaque état-membre est donc tenu de maîtriser l'incertitude de l'évaluation, mais c'est à lui qu'il revient de fixer le niveau à atteindre et les modalités d'estimation à utiliser.

La DCE ne donne pas non plus de précision en ce qui concerne la méthode statistique à appliquer. Néanmoins, une proposition d'utilisation de tests statistiques de « randomisation » (probabilité d'attribuer une classe erronée à un site) est faite.

Afin de comprendre l'étendue de la diversité des incertitudes en jeu au cours du protocole IBMR, il faut connaître les différentes étapes du protocole.
En effet, chaque étape de la méthodologie comporte une part d'incertitude, liée à la mise en œuvre des protocoles ou intrinsèque à la démarche. Le protocole d'évaluation IBMR comprenant toute une chaîne méthodologique, les incertitudes vont s'accumuler tout au long de cette chaîne.
Dans ce contexte, l'estimation des incertitudes en hydrobiologie doit être intégrée dans les règles d'évaluation utilisées pour l'état des lieux, le rapportage et le suivi de l'état écologique.

La Figure 4 représente les différentes étapes de la chaîne de surveillance/évaluation :

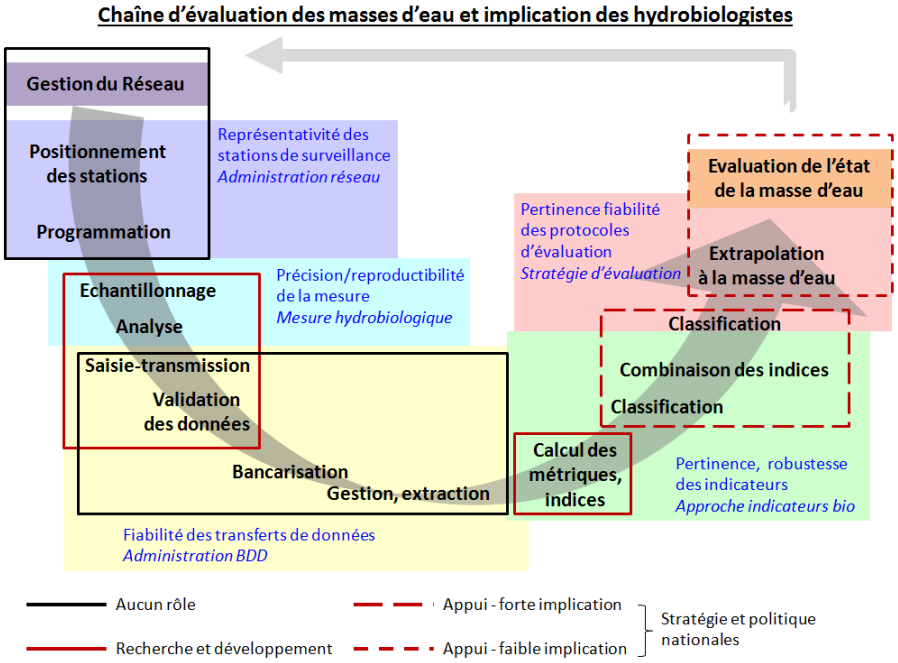


Figure 4 : Chaîne d'évaluation des masses d'eau et étapes impliquant les approches hydrobiologiques [6]

Ce qui nous intéresse ici pour notre travail est la partie « Echantillonnage – Analyse », ainsi que la partie « Calcul des métriques, indices - Classification ».

Echantillonnage – analyse [6]

Le protocole d'acquisition des données hydrobiologiques se décompose en deux étapes :
- l'échantillonnage, qui fait le lien entre le peuplement réel et l'idée que l'on s'en fait par l'analyse ;
- « l'analyse », qui consiste à traiter, identifier et quantifier le matériel biologique prélevé pour établir une liste taxinomique.

La « mesure hydrobiologique » finale obtenue comprend une/des liste(s) taxinomique(s) avec l'abondance de chaque taxon. L'incertitude sur cette mesure peut être évaluée en étudiant séparément les sources d'incertitudes de chaque étape du protocole d'acquisition de la donnée, et celles existant lorsque elles se combinent.

Calcul des métriques, indices - classification - combinaison des indices [6]

Le calcul de la valeur de l'indice permet ensuite l'attribution d'une classe d'état écologique au site évalué selon l'élément de qualité biologique traité. La confiance sur l'attribution de la classe d'état est directement influencée par les choix qui ont présidé à la définition des bornes de ces classes.
Pour ce même site, plusieurs indices sont combinés afin d'obtenir une seule classe de qualité finale. Une incertitude sur cette combinaison d'indice existe également.

Schéma conceptuel [6]

Le schéma (cf figure 5) représente une part importante de l'incertitude de la méthode d'évaluation, elle se divise en 4 parties:
- la variabilité du milieu évalué et des biocénoses présentes, (1)
- le protocole d'acquisition des données hydrobiologiques, (2)
- le calcul des indicateurs biologiques (notamment l'IBMR), (3) et
- la classification de l'état écologique. (4)

Les sources d'incertitudes sont détaillées au sein de chaque partie.

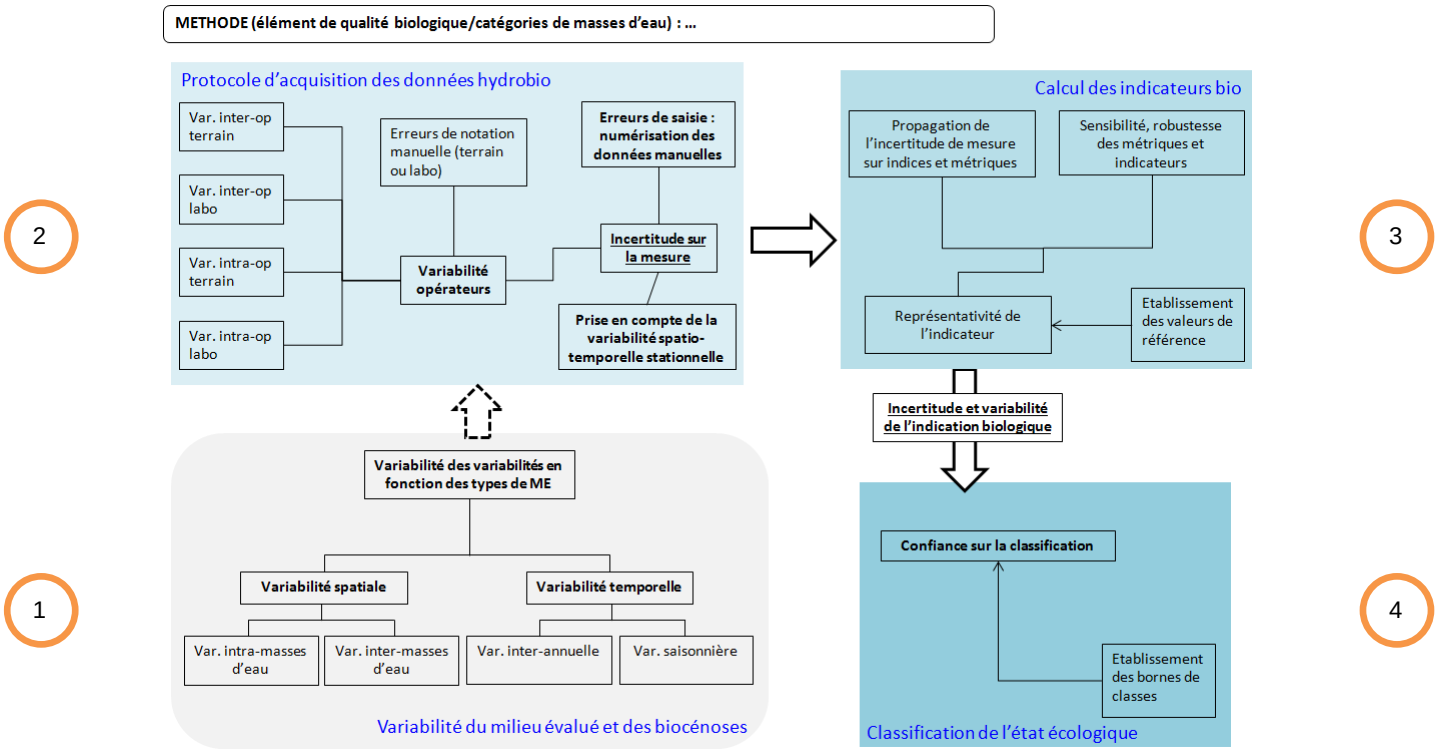


Figure 5 : Schéma conceptuel de l'identification des incertitudes dans le processus d'évaluation des masses d'eau [6]

L'incertitude va se présenter sous différentes formes pour chacune de ces étapes d'évaluation de la qualité des eaux.

L'incertitude sur la mesure est constituée de nombreuses sources [6] et va se propager sur les métriques et indices (dont l'IBMR), puisque c'est cette mesure qui est utilisée pour le calcul de ces derniers.
Les métriques et indices sont eux-mêmes sujets à une incertitude au niveau de leur construction, qui, associée à l'incertitude sur l'indice issue de l'acquisition de la donnée, représente une part de l'incertitude de l'indication biologique.

Enfin, le passage de l'indication biologique à la classification de l'état écologique intègre une incertitude représentée par la confiance sur la classification.

b. L'incertitude dans la littérature scientifique

Selon la littérature, les différentes sources d'incertitudes sur le résultat de l'évaluation écologique d'une masse d'eau peuvent se répartir en trois voire quatre catégories : [6] [23]

- La variabilité spatiale naturelle (sur un site de mesure, entre sites d'un tronçon)
- La variabilité temporelle naturelle (interannuelle, annuelle, saisonnière, mensuelle, jour/nuit)
- La variabilité liée à l'opérateur : positionnement du point de mesure, erreurs d'échantillonnage, d'analyse biologique, d'identification. Ce sont les erreurs de mesure. Contrairement aux autres sources de variabilité, celle-ci peut être partiellement contrôlée.
- La variabilité spatiotemporelle, qui constitue donc une interaction entre les 2 premières, a lieu quand un effet temporel agit différemment selon la location.

Le risque de mauvaise classification de la qualité des eaux est abordé par la DCE, avec les probabilités de mauvaise classification de ces masses d'eau.
La DCE met en avant le risque de classification erronée plus important au niveau des limites de classe de qualité qu'au milieu de ces classes.

Ceci est notamment constaté dans les articles [23], [24], [6] qui présentent des graphiques tels que les suivants, où l'on peut voir la nette différence du risque de mauvaise classification en fonction de la proximité de l'EQR des masses d'eaux avec les limites de classe :

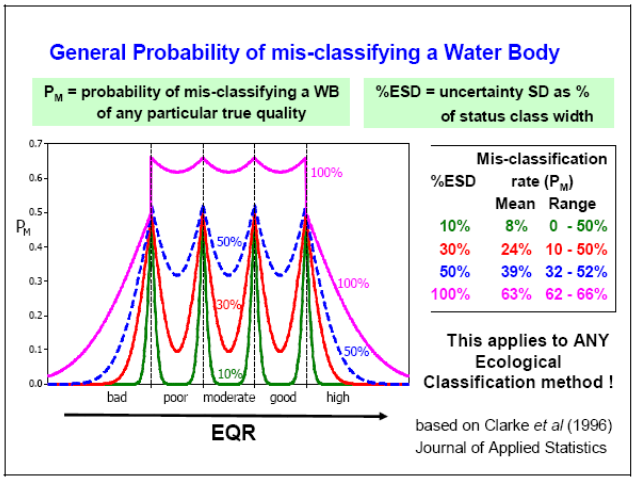


Figure 6 : Expression des probabilités de classement erroné dans les classes de qualité (d'après Clarke *et al*., 1996) [22]

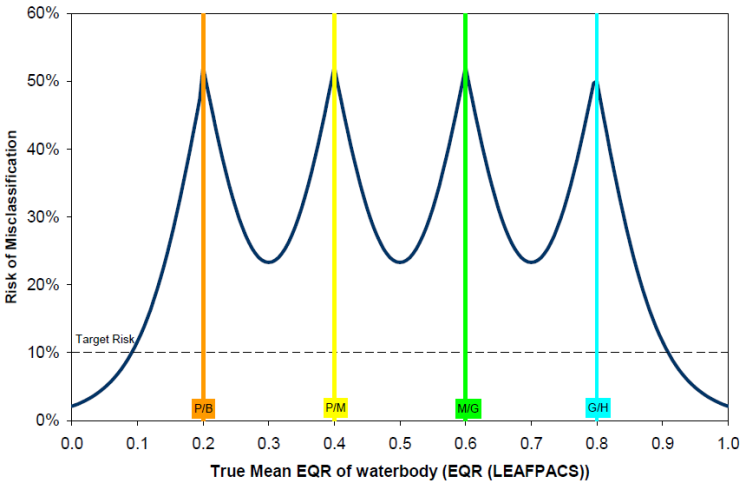


Figure 7 : Risque de mauvaise classification des eaux pour chaque classe de qualité selon les différents types de variabilité [23]

Les études sur les incertitudes liées aux populations macrophytiques se révèlent relativement peu nombreuses.

Quelques travaux ont permis d'obtenir des informations concernant la variabilité inter-opérateurs ou temporelle de plusieurs métriques et indices d'évaluation « macrophytes en cours d'eau » ainsi que d'évaluer la sensibilité de ces métriques et indices. Cependant, le nombre d'études étant limité, il n'est pas possible de conclure sur l'importance ou la quantification de ces différents types d'incertitudes.
De cette façon, il n'est pas encore possible de satisfaire la demande de la DCE sur l'évaluation de la confiance et de la précision des méthodes et de la classification intégrant toutes les sources d'incertitude. [6]

Le bilan des connaissances et informations disponibles à l'heure actuelle pour renseigner les différents domaines de l'incertitude se résume comme ci-dessous :

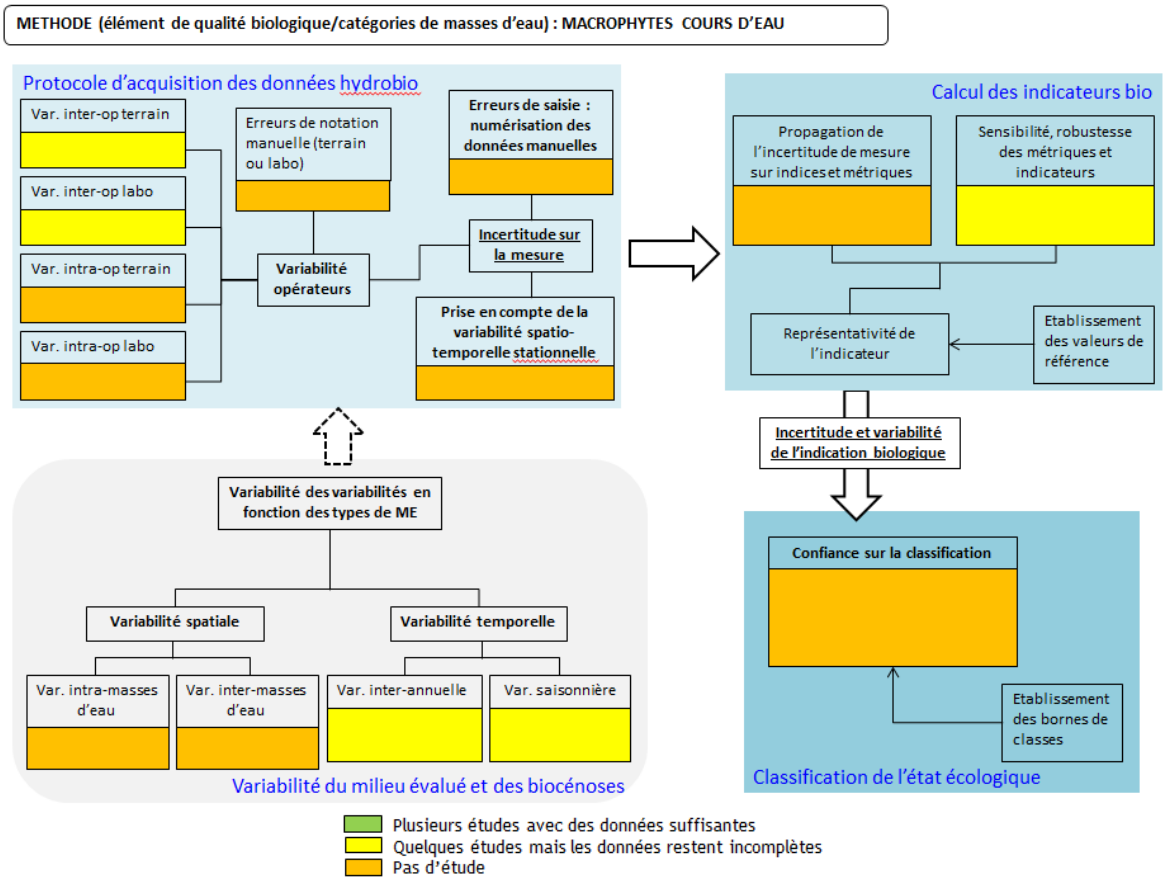


Figure 8 : Bilan de la littérature sur l'incertitude des "macrophytes en cours d'eau" [6]

Ces différentes études ont permis d'évaluer les différentes variabilités existantes lors de l'application de la méthode « macrophytes en cours d'eau » bien que ce soit principalement centré sur l'application du protocole MTR (« Mean Trophic Rank, indice macrophytique anglosaxon). Cependant, d'après Pentecost *et al.* 2009 [26], les méthodes d'acquisition des données et d'évaluation diffèrant peu entre pays européens, les protocoles et indices MTR et IBMR peuvent en conséquence être estimés comme similaires.. On peut alors utiliser les résultats obtenus dans ces différentes études afin d'évaluer l'incertitude liée au protocole d'acquisition de la donnée IBMR. [6]

Une étude a été lancée début 2013 sur le protocole et l'indice « IBMR » : Estimation des incertitudes liées à l'acquisition des données « macrophytes » en cours d'eau. Cette étude vise à estimer les variabilités suivantes [6]:

- Variabilité liée à l'opérateur : analyse des données obtenues par 3 opérateurs différents.
- Variabilité liée au positionnement du point de prélèvement : relevés successifs de 50 m sur 500 m au total et analyse des données obtenues.
- Variabilité saisonnière : 3 relevés pendant une période de végétation.
- Variabilité interannuelle : suivi de plusieurs stations sur 6 ans.
- Variabilité des métriques et indices liées aux erreurs de détermination (le sujet de ce rapport)

Les résultats seront disponibles pour une grande partie début 2016.

Les incertitudes en hydrobiologie peuvent ainsi être étudiées à différents niveaux : sur la mesure (liste taxinomique et abondance des taxons), sur les métriques, sur les indices, sur les classes d'état, etc. De nombreux travaux à ce sujet sont donc à prévoir.

c. Les travaux de J. Wiederkehr

L'incertitude inter-opérateur liée à la détermination taxinomique n'a, à ce jour, été traitée qu'une seule fois, et récemment par J.Wiederkehr [27] [28]

L'étude de J. Wiederkehr se concentre sur l'effet des erreurs d'identification taxinomique (effet opérateur) sur l'IBMR et le niveau trophique.
Pour ce faire, ont été créées des modifications virtuelles aléatoires des taxons identifiés et des classes de recouvrement associées, en se basant sur 2 matrices de confusion et un algorithme spécifique.
L'incertitude est représentée dans la matrice de confusion par des coefficients de confusion dont les valeurs ont été attribuées par des experts selon leur expérience sur le terrain et en laboratoire.

L'algorithme prend comme données d'entrée une liste floristique et un taux d'erreur (coefficients de confusion), et produit une liste de taxons ou classes de recouvrement modifiés de façon aléatoire en fonction de ce qui a été rentré : soit la matrice de confusion des taxons, soit la matrice de confusion des classes de recouvrement.
Par la création de ces nouvelles listes floristiques obtenues à partir d'erreurs définies de façon aléatoire et par le calcul des métriques résultantes, les effets de ces erreurs sur l'IBMR et les statuts trophiques ont pu être déterminés.

3 ensembles de tests ont été développés pour chaque type de liste floristique vérifiée :

(1) Un test avec seulement les modifications de taxon (TAX)
(2) Un test avec les modifications de classe de recouvrement (AC)
(3) Un test avec ces deux types de modification (TAX + AC)

Grâce à ces méthodes, ils ont pu déterminer que :

- les confusions de taxons (TAX) et les deux erreurs combinées (TAX +AC) ont toujours un effet plus important que les confusions sur les classes d'abondance seules (AC)
- il est plus probable qu'il y ait changement de classe quand le score est proche des limites de classe.
- le déclassement pour un cours d'eau est plus probable que le surclassement
- les courants de niveaux trophiques faibles et moyens ont une probabilité plus élevée d'être déclassés que les autres, ce qui est probablement lié au petit intervalle de ces classes.
- il y a moins d'erreurs significatives pour les petits taux d'erreur (10 et 25%) que pour les forts taux d'erreur (50, 75, et 100%).

Finalement, J. Wiederkehr fait des propositions pour limiter l'impact des incertitudes et obtenir des résultats plus fiables, comme entraîner des opérateurs, faire des confrontations *in situ* des avis entre opérateurs, et appliquer une approche de contrôle qualité.
Il apparait évident qu'il devient nécessaire de faire spécialement attention aux taxons pouvant être confondus et dont les coefficients spécifiques CS sont éloignés (ce qui augmenterait l'impact sur le score IBMR).

Cette méthodologie reposant sur l'utilisation de matrices de confusion est comparable à l'approche suivie dans le programme Onema et Irstea, présentée dans ce rapport. Les aspects qui diffèrent significativement sont principalement l'échelle géographique (régionale vs. nationale), ou encore la quantité des taxons considérés dans les confusions potentielles.

3) Résultats obtenus

Un répertoire rassemblant tous les couples de confusion répertoriés dans les différentes sources à disposition (guides de détermination) a été conçu. Ce fichier constitue une analyse bibliographique préliminaire et sera aussi utile pour les différents botanistes intéressés par les risques de confusion. Il regroupe à la fois les couples de taxons pouvant porter à confusion, classés par groupe taxinomique, mais aussi des remarques sur ces confusions, et la/les source(s) utilisées indiquant la possibilité de cette confusion. Ce fichier correspond donc à un croisement des données de différentes sources.

Dans un même temps, il a fallu envoyer aux experts participant à la mission, un fichier de tableaux rassemblant des couples de taxons, préparé antérieurement par l'équipe, ainsi qu'une notice explicative de la démarche à entreprendre, et une liste des différentes espèces avec leurs synonymes associés.
Il y eu par la suite une synthèse de ces fichiers (un par expert), pour produire une matrice se basant sur l'ensemble des avis de ces différents experts quant aux classes de confusion des différents couples.

Cette matrice, appelée matrice de confusion et d'une taille de 460x460, rassemble l'ensemble des coefficients de confusions estimés pour les différents couples étudiés, et se présente sous un format utilisable pour les analyses numériques.
Elle a permis de créer par la suite des listes floristiques dégradées (2.629.000 au total) en se basant sur des listes floristiques présentes sur la base de données Pandore. L'IBMR obtenu à partir de ces listes dégradées a alors pu être comparé à l'IBMR des listes non dégradées afin de comparer l'effet de l'incertitude liée à l'identification taxinomique sur le score de l'IBMR.

Cette matrice de confusion, élément clef de ce stage, est donc une étape primordiale et nécessaire afin de répondre aux exigences de la DCE quant à la «confiance» et la «précision».

Des tests statistiques ont alors été faits sur ces différents résultats afin d'analyser l'impact des confusions taxinomiques sur l'indicateur IBMR et par la suite sur les classes de qualité. A ce sujet, une certification MOOC sur R a été obtenue. [29]

En outre, une participation à la première campagne sur le terrain du programme « variabilité stationnelle » a été assurée, afin de concrétiser les connaissances acquises sur le protocole de relevé des macrophytes en cours d'eau et de visualiser certaines sources d'incertitude de mesure, comme la répartition des peuplements et des espèces dans un tronçon de cours d'eau présenté comme homogène.

II- Outils et démarche scientifique

Le travail de stage réalisé a été mené en interaction étroite avec les ingénieurs de l'équipe Irstea, qui sont chacun impliqués dans ce programme de recherche et développement sur les incertitudes « méthode macrophytes cours d'eau ».
La méthode et les protocoles suivis ont été définis au regard des objectifs fixés dans le programme, et en fonction de l'avancée des analyses. Pour ce travail de stage, le programme a été ajusté et redéfini en fonction du temps pris effectivement par chacune des phases et du temps contraint pour le stage.

1) La base de données PANDORE

La base de données PANDORE [30] a été mise en place par Irstea dès 2005, afin de regrouper sous un format homogène les données issues des réseaux de surveillance de la qualité des eaux de rivière disponibles dans chacune des Agences de l'eau. Irstea a ensuite assuré la collecte et la bancarisation des données hydrobiologiques générées par la mise en œuvre des nouveaux réseaux de surveillance en application de la DCE.

Physiquement hébergées sur les serveurs du centre Irstea de Lyon, cette base sous moteur SQL peut être manipulée depuis le centre de Bordeaux via des protocoles informatiques sécurisés.

PANDORE contient donc les répertoires des stations de mesure sur tout le réseau hydrographique français (plus de 4000 stations à ce jour), des données hydrobiologiques sur les macrophytes, les diatomées et les invertébrés, ainsi que les données physico-chimiques.
PANDORE est donc une base de données complète de portée nationale, sur laquelle s'est appuyée l'élaboration des méthodes nationales d'évaluation d'état écologique par les équipes d'Irstea.
Elle constitue la matière première de toutes les approches menées dans les programmes de recherche et développement actuels sur les bioindicateurs et les méthodes d'évaluation.

Cette base de données est celle sur laquelle reposent les travaux d'analyse des incertitudes, dans lesquels s'inscrit ce travail de stage.

2) Fichier répertoire des couples sujets à confusion

Ce fichier constitue une analyse bibliographique préliminaire, permettant de se faire une idée des confusions possibles chez les macrophytes.

La conception de ce fichier constitue un travail en soi, et permet d'identifier des sources bibliographiques faisant référence aux confusions taxinomiques.
Cette approche présente de ce fait le double objectif d'inventorier les espèces les plus sujettes à confusion, et de constituer un répertoire regroupant des informations issues de différentes sources, susceptible de servir de base aux choix lors du traitement de l'avis des experts dans la synthèse de la matrice de confusion, présentée dans la suite du rapport.

Ce fichier a été constitué sur la base d'un certain nombre d'ouvrages de détermination pratique présents dans les locaux d'Irstea (car utilisés fréquemment par les hydrobiologistes), qui mentionnent les risques de confusion possible dans une approche de détermination floristique pratique. Il pourra être complété avec d'autres ouvrages, ou quand de nouvelles confusions seront identifiées.

Pour élaborer ce fichier répertoire, 9 ouvrages ont été sélectionnés (cf Annexe 2). [31-39] Parmi ces 9 ouvrages, certains étaient spécifiques à un groupe taxinomique (Algues, Characées, etc d'autres traitaient de façon générale de différents groupes (« végétaux aquatiques », etc.).

Le fichier a été créé sous format Excel, avec un onglet pour 5 des 7 groupes taxinomiques considérés dans les « macrophytes » (cf tableau 4) :

- Algues
- Bryophytes,
- Characées
- Phanérogames
- Ptéridophytes

Pour chaque taxon identifié comme sujet à confusion, on obtient le résultat suivant :

Tableau 4 : Extrait d'un tableau du répertoire de confusion taxinomique

Confusion possible			Sources bibliographiques		
Taxon A	Taxon B	Remarques	(Fare et al., 2011)	(Mouronval et al., 2010)	Spencer-Jones et al., 1986
Althenia filiformis	Ruppie maritime			X	
Althenia filiformis	Zannichellia pedunculata			X	
Althenia filiformis	Najas gracillima			X	
Apium inondatum L. *Apiaceae*	Renoncules aquatiques	En l'absence de fleur	X		

Dans ce tableau, « Taxon A » est le taxon de référence, « Taxon B » liste ceux avec lesquels il peut être confondu. Les remarques potentiellement explicatives des confusions ou des conditions dans lesquelles elle peut advenir sont consignées dans « Remarques ». Les sources bibliographiques sont précisées par la présence ou l'absence d'un « X ».

Un dernier onglet rassemble les références bibliographiques précises, par groupe taxinomique. (cf Annexe 2)

Pour compiler ces informations de façon homogène, les problèmes d'appellation taxinomiques ont été traités : différences de nom vernaculaire, synonymies, nom scientifique accompagné ou non du nom de la famille, nom d'auteur, etc. Après vérifications, seuls les binômes (genre et espèce) ont été conservés.

A partir de ce fichier, un fichier utilisable numériquement pour les analyses, en particulier sous l'application R, a été créé, en ne conservant que les informations utiles à cette utilisation :

- Suppression des colonnes rassemblant les sources bibliographiques
- Ajout des codes « 6 lettres » (6L) (cf tableau 5)
- Contrôle de cohérence et de complétude des appellations et des codes 6L

Nota : Les codes 6L sont issus du référentiel taxinomique mis en place et géré par Irstea. Ce référentiel est utilisé dans les applications et les bases de données pour permettre le calcul des indicateurs à partir des données issues de listes floristiques, en minimisant les risques d'erreurs liés aux problèmes de retranscriptions ou de graphie des noms scientifiques complets.

Tableau 5: Onglet contenant la liste taxinomique de référence avec les codes 6L

nom_taxon_ListFlore	cd_taxon
Achillea millefolium	ACHMIL
Achillea ptarmica	ACHPTA
Achnatherum calamagrostis	ACNCAL

Les listes de confusions ont fait l'objet d'une validation par les experts de l'équipe Irstea, afin de les adapter à l'utilisation prévue et au contexte des données (suppression de certains couples inutiles, ajout d'autres non mentionnés dans les guides de détermination consultés, création de codes 6L non pré-existants, etc.).

Cette étape a donc permis de produire :

- Un fichier Excel comportant les sources bibliographiques, les taxons avec les différentes précisions (sous-espèces, etc.), les noms vernaculaires, les remarques et informations utiles ;
- Un fichier Excel ne contenant que les informations brutes sous un format utilisable facilement par les applications informatiques telles que R.

Ce répertoire constitue un élément nouveau et pourra être utile lors de la détermination taxinomique des macrophytes en identifiant les différentes confusions possibles.

3) Matrice de confusion

a. Principe et contenu de la matrice de confusion

La matrice de confusion constitue l'outil de base pour le calcul du jeu de listes floristiques dégradées (cf II.4 « La dégradation des listes »), qui vont permettre l'analyse de la propagation des erreurs dans la chaine d'évaluation de l'état écologique des cours d'eau.
Il s'agit avec cette matrice de confusion d'estimer les erreurs susceptibles d'être faites au niveau de la détermination taxinomique par les opérateurs, lors de relevés de macrophytes réalisés dans le cadre de la surveillance DCE selon le protocole décrit dans la norme NF T90-395 [11].

La conception de cette matrice a compris plusieurs étapes principales :

- Formalisation d'une matrice complète des taxons PANDORE ;
- Présélection d'une liste de couples de taxons susceptibles de présenter des risques de confusion ;
- Constitution d'un fichier de consultation des experts, permettant un recueil d'informations claires et homogènes ;
- Enquête auprès d'un panel d'experts ;
- Compilation, harmonisation et synthèse des réponses ;
- Constitution d'un fichier utilisable dans les traitements informatiques prévus.

Le principe de la matrice est de préciser un risque de confusion, sous la forme d'un coefficient de confusion, pour chacun des couples de taxons.

Tableau 6 : Schéma du format final de la matrice de confusion

	Taxon 1	Taxon 2	Taxon 3
Taxon 1	a	b	c
Taxon 2	b	d	e
Taxon 3	c	e	f

La diagonale représente la non-confusion (bonne détermination).
Ainsi on peut voir que le taxon 1 a un coefficient b de confusion avec le taxon 2, un coefficient c avec le taxon 3, etc.

b. Constitution du fichier de consultation des experts

Afin de faciliter le travail des experts et d'expliciter la demande, pour obtenir des réponses plus rapides et plus homogènes, un fichier de consultation des experts a été élaboré. Il comprenait, sous un format de classeur Excel :

- Une feuille pour chaque groupe taxonomique (7 onglets au total) ;
- Pour chaque feuille, une liste de couples de taxons susceptibles de présenter des risques de confusion ;
- Une feuille d'explication et de consignes.

<u>Sélection des couples de taxons</u>

Une matrice complète a d'abord été générée en créant tous les couples possibles à partir de la liste des taxons présents dans la base de données PANDORE.

Cette liste a été filtrée selon deux critères :

- Le coefficient d'aquaticité des taxons, afin de ne retenir que ceux directement liés au milieu aquatique du cours d'eau, en excluant les espèces de berges ou rudérales. Cette aquaticité (attribut de chaque taxon figurant dans PANDORE) est définie selon le tableau 7. Seuls les taxons d'aquaticité 1, 2 ou 4 sont retenus ;
  Cette première approche a généré une matrice de 351 taxons.

Tableau 7 : Définition des groupes d'aquaticité *(en grisé, les groupes strictement inféodés à l'eau)*[40]

Aquaticité	Affinité pour l'eau
1	Taxons exclusivement ou principalement aquatiques en conditions normales d'étiage
2	Taxons aquatiques présentant communément une forme terrestre et une forme aquatique, ou régulièrement amphibies
3	Bryophytes et lichens supra-aquatiques (peuvent être immergés une partie de l'année)
4	Hélophytes ou amphiphytes, normalement rencontrés le pied dans l'eau
5	Taxons hygrophiles, susceptibles d'être rencontrés le pied dans l'eau
6	Taxons herbacés rivulaires, forestiers, prairiaux ou rudéraux, accidentellement en position aquatique
7	Arbustes et arbres de la ripisylve (ligneux), pouvant être temporairement inondés
8	Taxons des eaux saumâtres et marais salants

Nota : Les coefficients d'aquaticité, caractérisent l'affinité pour l'eau des taxons pris en compte dans les approches « macrophytes ». Ces coefficients ont été proposés par Irstea pour l'élaboration de la base de données européenne « macrophytes cours d'eau » dans le cadre des travaux d'intercalibration des méthodes européennes [40].

- L'élimination à dire d'expert des couples ne présentant notoirement aucun risque de confusion. Cette sélection ne conservant que les couples présentant un risque potentiel, au moins dans certains cas, permet un travail plus simple pour les experts sollicités. Ce travail a été réalisé par les experts de l'équipe Irstea.

Ces couples de taxons, listés dans un fichier Excel, ont été regroupés par groupes taxinomiques :

- Algues
- Bryophytes
- Characées
- Hétérotrophes
- Lichens
- Phanérogames
- Ptéridophytes

Chaque groupe fait l'objet d'une feuille dans le fichier Excel.

Un onglet pour les confusions inter-groupes éventuelles a été ajouté.
Un dernier onglet à renseigner a été expressément établi afin que les experts puissent y ajouter les couples supplémentaires qu'ils jugeraient important de noter par rapport à leur confusion possible, qui n'auraient pas été listés dans le fichier proposé.

Enfin, dans le fichier Excel a été ajouté un onglet avec la liste de référence des noms des espèces/genres et leurs synonymes connus, afin de guider les experts pour la recherche éventuelle des appellations.

Un mode de renseignement des listes a été proposé, en utilisant, pour chacun des couples que l'expert choisi de renseigner, 3 classes de confusion :

- 0 - pas de confusion possible,
- 1 - confusion peu fréquente, possible pour certaines formes ou par des opérateurs peu familiers de ces groupes,
- 2 - confusion assez courante dans certaines conditions, pour certaines formes ou pour des opérateurs non spécialistes de ces groupes/genres,
- 3 - confusion très fréquente dans la plupart des cas, pour toutes les formes du taxon, pour tous les opérateurs non spécialistes de ces groupes ou des types de peuplements concernés.

Chaque expert peut ne renseigner que certains groupes, ou renseigner uniquement les taxons sur lesquels il a un avis.

Il a été demandé aux experts de tenir compte de l'asymétrie éventuelle des confusions. En effet pour un même couple, la probabilité de confusion n'est pas forcément la même dans les 2 sens (probabilité de confondre A avec B ≠ probabilité de confondre B avec A), par exemple entre une espèce A ubiquiste à très large répartition et une autre B spécifique de certaines régions ou types particuliers de cours d'eau (B sera alors souvent dénommée A, alors que l'inverse sera très rare). Les experts ont donc eu la possibilité de noter cette particularité.

Une des consignes importantes est d'évaluer ces confusions en considérant le cas d'opérateurs ayant la compétence et l'expérience dans l'identification d'espèces nécessaire et suffisante pour la mise en œuvre du protocole décrit dans la norme, mais n'étant pas spécialistes des différents groupes de macrophytes. En effet, les opérateurs des réseaux de surveillance sont généralement des hydrobiologistes de bureaux d'étude, dont les compétences sont le plus souvent suffisantes pour l'application en routine, mais non spécialisée en floristique. Cette identification est alors faite en utilisant des guides ou des ouvrages courants, et, le cas échéant, du matériel d'optique standard dans un laboratoire d'hydrobiologie, dans un cadre de prestation de service standardisée. De même, dans la conduite d'un programme de relevé sous forte contrainte de planning, les conditions d'observation peuvent ne pas être optimales (relevé hors période de développement maximal ou hors période de floraison, par exemple). L'évaluation des risques d'erreurs de détermination doit tenir compte de ce contexte.

Le référentiel utilisé pour ce travail est celui d'Irstea pour les appellations de taxons, fourni dans le premier onglet du document Excel à remplir. Ceci est important, car les nombreuses synonymies amènent très souvent un biais dans le travail sur des listes de taxons par des experts différents, qui utilisent parfois des appellations différentes.

Une notice à l'attention des experts a été élaborée (cf Annexe 3), qui reprend l'ensemble des points explicités ci-dessus : structure du classeur, informations attendues et consignes. Cette notice est jointe au fichier Excel.

c. Consultation des experts

Dans l'objectif de construire une matrice de confusion complète et regroupant un consensus suffisant entre expert pour assurer une robustesse satisfaisante à cette approche, un panel d'experts a été sollicité.
Ces experts sont des botanistes ou phytoécologues impliqués dans la détermination de macrophytes et connaissant le protocole de l'IBMR et le contexte DCE. Ces experts ont également été choisis pour leur activité de validation taxinomique d'échantillons récoltés par les bureaux d'étude prestataires des réseaux DCE, ce qui permet de bénéficier d'un retour d'expérience large sur les confusions régulièrement commises.

Le panel d'experts a principalement été sollicité dans le cadre du réseau de macrophytistes du GIS (groupement d'intérêt scientifique) « macrophytes des eaux continentales ».

Le GIS « macrophytes des eaux continentales » est un réseau de compétence de scientifiques créé dans le but de « structurer la recherche et la réponse aux gestionnaires dans le domaine des plantes aquatiques ». L'équipe d'Irstea Bordeaux fait partie de ce GIS depuis sa création en 1992. [41]

Ses activités portent sur les communautés végétales des cours d'eau [41]:

- Mise en place d'un protocole standardisé d'analyse de la végétation des cours d'eau
- Elaboration puis proposition d'indice de qualité biologique des cours d'eau _ indices GIS
- Synthèse sur les méthodes de bioindication puis normalisation de l'IBMR et développement de cet indice.

La démarche de cette étude a été présentée aux membres du GIS lors de ses journées annuelles en mars 2015 par les chercheurs de l'équipe Irstea Bordeaux.

Au final, le panel d'experts ayant apporté leur contribution à cet exercice a été constitué de :

- Monique Aubert, Hydrobiologiste consultante indépendante (tous les groupes)
- Marie Christine Peltre, Université de Lorraine-Metz (algues)
- Christophe Laplace-Treyture, Irstea Bordeaux(algues)
- Julie Coudreuse, Agrocampus Ouest de Rennes (bryophytes)
- Elisabeth Lambert, Université Catholique de l'Ouest - Angers (characées)
- Jacques Haury, INRA- Agrocampus Ouest de Rennes (phanérogames)
- Christian Chauvin Irstea Bordeaux (tous les groupes)
- Sandrine Loriot Irstea Bordeaux (tous les groupes)

Ainsi ont été envoyés à chaque expert, le 8 avril 2015 :

- un mail rappelant brièvement l'intérêt de cette matrice -ce qui avait été fait au préalable à l'oral lors d'une réunion- et indiquant les différents fichiers joints à ce mail
- le fichier Excel à remplir
- la notice explicative

d. Homogénéisation des réponses des experts

Un script sous R a été élaboré pour permettre une compilation automatique et un contrôle de conformité des informations retournées par les experts.

Ce script a été testé sur les premiers fichiers collectés, pour achever sa mise au point par l'élimination des anomalies apparaissant à l'ajout de chacun de ces fichiers. Le script a de cette façon pu être perfectionné pas à pas. Des modifications ont de même été apportées directement, soit manuellement, sur les fichiers Excel.

A chaque réception de document rempli par un des experts, il a fallu faire un travail de mise en forme et d'homogénéisation afin de rendre utilisables les documents avec le script R :

- Homogénéisation du nom des documents faisant apparaître le nom des experts en question ;
- Présence d'un total de 9 onglets correspondant aux 9 « groupes taxinomiques » précisés plus haut, et suppression de l'onglet contenant la liste de référence des synonymes ;
- Format sous 4 colonnes au sein de chaque feuille, avec pour noms : « nom_taxon_A », « nom_taxon_B », « classe_de_confusion (de 0 à 3)», « remarques » ;
- Homogénéisation de la police, suppression du surlignage, des couleurs, etc. ;
- Conservation des commentaires utilisables pour l'attribution des classes de confusion de taxons ;
- Attribution des codes 6L aux taxons A et B de chaque couple ;
- Contrôle de cohérence et de complétude des appellations et des codes 6L ;
- Correction des anomalies détectées dans les appellations (orthographe, ajout/manque de points/d'espaces, niveaux de détermination ne correspondant pas aux exigences de la méthode [11] (au genre pour les unicellulaires comme les algues et les hétérotrophes, et de l'espèce pour les autres taxons), noms obsolètes, synonymies, appellations autres que les noms scientifiques, etc.) ;
- Suppression des doublons.

Les données de confusion considérées dans l'approche menée par Wiederkehr 2015 [27, 28] ont également été considérées, en tant qu'avis d'expert

e. Synthèse des avis des experts et création de la matrice

Une fois la vérification et les modifications des documents Excel et du script faites, le script a pu être définitivement lancé afin d'obtenir la synthèse de l'ensemble des informations apportées par les experts.

Le script a été conçu de manière générique afin de permettre à l'ensemble des matrices, y compris celles venant d'être rajoutées, d'être traitées automatiquement sans avoir besoin de modifier tout le script.

Pour rappel, les 9 onglets des 8 documents Excel des experts se présentent sous la forme suivante :

Taxon A	Taxon B	Classe confusion	Remarque

Figure 9 : Schéma de la structure d'un onglet du fichier de consultation des experts

Le but est d'obtenir un seul et unique document rassemblant toutes ces données.
Au final, 2 types de synthèse ont été produites :

- 1 document composé de 9 onglets (1 par groupe taxinomique), avec dans chaque onglet l'avis de tous les experts. Ce document a notamment été utile pour les tests statistiques.
- 1 document composé d'1 seul onglet comportant l'ensemble des groupes taxinomiques tous rassemblés, avec de même l'avis de tous les experts. Ce document a été obtenu à partir du document précédent et a pu être utilisé pour la conception de la matrice de confusion.

Le schéma ci-dessous représente le rassemblement/la synthèse du même onglet n de tous les documents d'experts.

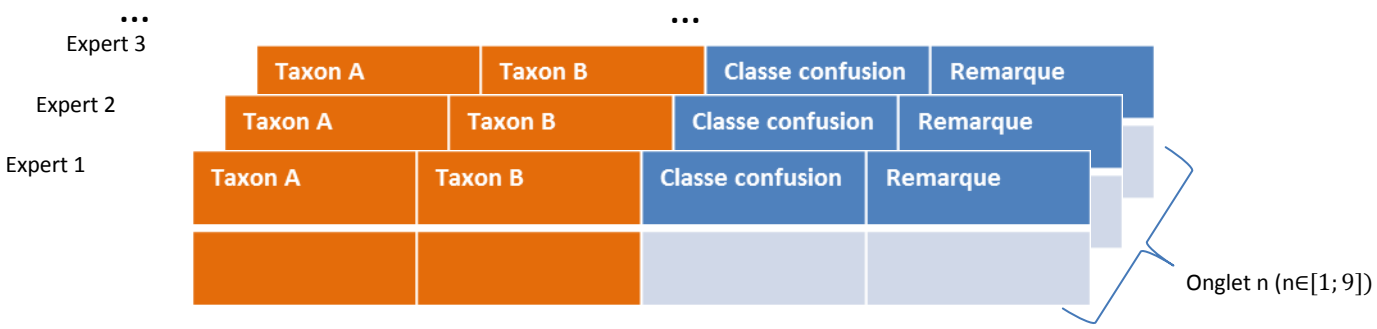


Figure 10 : Schéma de l'assemblage d'un même onglet de chaque fichier expert

Cet assemblage s'est fait en fonction des noms des taxons A et B qui constituent le lien de conjonction entre ces différents tableaux (en orange sur le schéma ci-dessus). De cette façon, s'il y avait présence d'un couple a et b dans plusieurs documents d'experts, alors l'ensemble des avis (en bleu sur le schéma ci-dessus) était assemblé sur la même ligne correspondant à la ligne propre au couple a-b.

Taxon A	Taxon B	Classe confusion	Remarque	Classe confusion	Remarque	Classe confusion	Remarque

Figure 11 : Schéma de l'assemblage des avis et remarques d'un couple

Des colonnes prévues pour l'avis et les remarques de chaque expert ont été créées afin que les classes de confusions et remarques s'ajoutent dans la colonne appropriée, à la ligne du couple en question. Cette opération s'est faite pour chaque onglet.

Pour chaque onglet final (pour chacun des 2 documents de synthèse), la mise en forme adoptée est la suivante :

Couple A_B	Taxon A	Taxon B	Classe confusion *Expert 1*	Remarque *Expert 1*	…	Classe confusion *Expert m*	Remarque *Expert m* (m: nombre d'experts)	Classe confusion définitive	Remarque définitive

« Commun » aux différents experts
⇨ Lien cohésif

Expert 1

Expert m

A établir selon l'avis des experts et le nôtre

Ajout classe de confusion + remarque de chaque expert pour chaque couple de cet onglet

Onglet n
(n ∈ [1;9])

Figure 12 : Schéma d'un onglet final, après synthèse des avis et remarques

La matrice de synthèse obtenue présente donc 9 onglets, chacun comportant un tableau de ce type.
Chacun de ces tableaux rassemble l'ensemble des couples appartenant à ce « groupe taxinomique » ainsi que leurs données associées : ensemble des avis des experts et leurs remarques.
Tous les couples de chaque onglet ont été triés par ordre alphabétique en fonction du nom du couple.
Une concaténation des avis a permis de faire apparaître dans chaque case d'une colonne conçue à cet effet, l'ensemble des avis des experts alors visualisables en un coup d'œil.

La prise en compte des informations issues des travaux de J. Wiederkehr (2015) [27, 28], a pu être assurée après vérification de la cohérence des codes 6L utilisés, le tri des taxons correspondant à des « macrophytes en cours d'eau » au sens de la définition retenue, ainsi que le filtrage des taxons ayant été observés au moins une fois dans la base Pandore.
Le format et le codage des probabilités de confusion n'étant pas compatibles avec la matrice constituée, les informations de J. Wiederkehr (2015) n'ont pas été intégrées directement dans la matrice, mais consultées en complément à la synthèse des avis d'experts.

Finalement, pour la colonne destinée à la classe de confusion définitive, ont été sélectionnés les maximums de l'ensemble des avis des experts. Le choix a donc été fait de considérer un risque d'erreur maximisé. .

Une des problématiques de la construction de cette première matrice de confusion fut le choix d'omettre volontairement l'asymétrie existant dans la confusion au sein des couples. En effet, comme expliqué dans la partie II.2.b. « Constitution du fichier de consultation des experts», la confusion de A vers B n'est pas forcément la même que de B vers A. Cette première approche est donc basée sur un concept de confusions symétriques, pour des raisons de complexité et de temps disponible pour la réalisation du présent travail.

Pour obtenir le document rassemblant les 9 onglets en un seul, le tableau de chaque onglet a été accolé à la suite des autres tableaux.

Le choix du scénario de recodage en % des codes de classe de confusion (cf II.5 « Aides à la décision dans la démarche ») est le suivant :
Confusion de classe 0 => 0% de risque de confusion
Confusion de classe 1 => 1% de risque de confusion
Confusion de classe 2 => 5% de risque de confusion
Confusion de classe 3 => 10% de risque de confusion

Il a aussi été nécessaire d'ajouter les taxons présents dans les cours d'eau mais qui avaient été supprimés du fichier de consultation des experts car non sujets à confusion. Pour cela une valeur de 100% leur a été attribuée pour la confusion avec eux-mêmes, et de 0% avec les autres taxons.

4) La dégradation des listes

Une fois la matrice de confusion créée, le protocole de dégradation des listes floristiques a pu être appliqué. Cette « dégradation » consiste en la génération aléatoire de confusions distribuées en cohérence avec les coefficients de la matrice, dans une liste floristique.

C'est par ces listes « dégradées », c'est-à-dire avec des erreurs volontairement induites, qu'il devient possible de comparer le nouvel IBMR « dégradé » avec l'IBMR calculé sur les listes floristiques originales, et donc d'évaluer les impacts provoqués par la confusion taxinomique sur l'IBMR.

Les valeurs d'IBMR « dégradés », les EQR, ainsi que les classes d'état correspondantes ont alors été calculés dans un premier temps à partir des listes floristiques originales, puis pour chacune des listes dégradées. Ces calculs ont été réalisés en appliquant un algorithme développé sous R.

On peut voir dans le schéma ci-dessous les 2 types d'approche, multiple et unitaire, utilisées pour étudier l'écart existant entre IBMR original et IBMR « dégradé ».

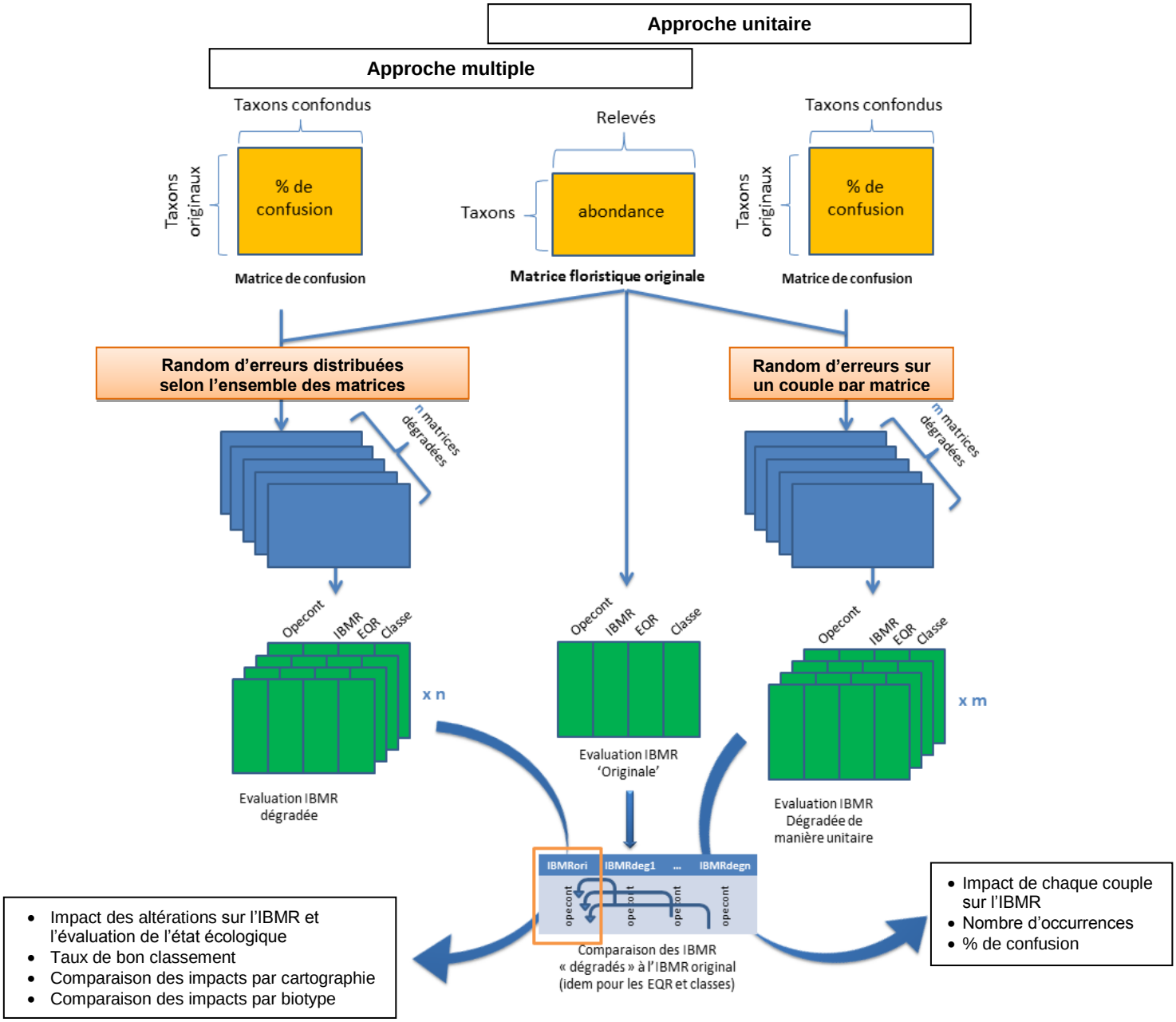


Figure 13 : Schéma de la démarche de dégradation des matrices et des résultats visés

Dans les 2 cas, il est question de comparer les résultats avec les valeurs issues de la matrice floristique originale (partie centrale du schéma). A partir de cette matrice ont été calculés les IBMR de chaque opération de contrôle, ainsi que les EQR et classes d'état issues de ces données.

L'approche unitaire (partie droite du schéma) consiste à ne dégrader qu'un seul couple pour chaque matrice dégradée. On obtient m matrices dégradées où chaque couple de confusion a été dégradé unitairement.
A partir de ces matrices dégradées ont été calculées des valeurs d'IBMR correspondants, Sur ce jeu de valeurs de l'indicateur, les classes d'état écologique ont ensuite été définies, en calculant, pour chaque valeur d'IBMR, l'écart à la référence puis en appliquant les seuils de classes d'état.

L'approche multiple (partie gauche du schéma) consiste à dégrader chaque matrice, en totalité, en fonction de la matrice de confusion et de ses coefficients de confusion attribués pour chaque couple. On obtient ainsi n matrices dégradées (n le nombre d'itérations du script, soit 1000 ici). A partir de ces matrices dégradées ont pu être calculés, tout comme l'approche unitaire, les IBMR dégradés correspondants, les EQR, et les classes d'état qui en découlent.

5) Aides à la décision dans la démarche

Des statistiques et autres analyses ont été utiles tout au long de la conception de la matrice de confusion, et l'ont de même été afin de comprendre les résultats obtenus.
En effet, à différentes étapes de la conception de la matrice de confusion, il a fallu faire des choix grâce à l'utilisation d'outils statistiques.

Des analyses ont tout d'abord été faites dans la synthèse des avis des experts dans le but d'éviter les anomalies liées à l'assemblage (doublons, suppressions, ajouts, etc.).

Comparaison des avis des experts

Les avis des experts pour chaque taxon ont par la suite été analysés afin de vérifier si les avis ne variaient pas trop pour chaque couple de confusion, ce qui serait signe d'une forte variabilité pouvant exister dans l'identification taxinomique en fonction de l'opérateur. La classe de confusion résultante en serait moins fiable.
La différence dans les avis d'experts, que ce soit par rapport à la quantité de couples traités ou la moyenne des classes attribuées, a été illustrée pour chaque groupe taxinomique afin d'évaluer la variabilité inter-experts. On peut voir par exemple ci-dessous les graphiques concernant les Bryophytes.

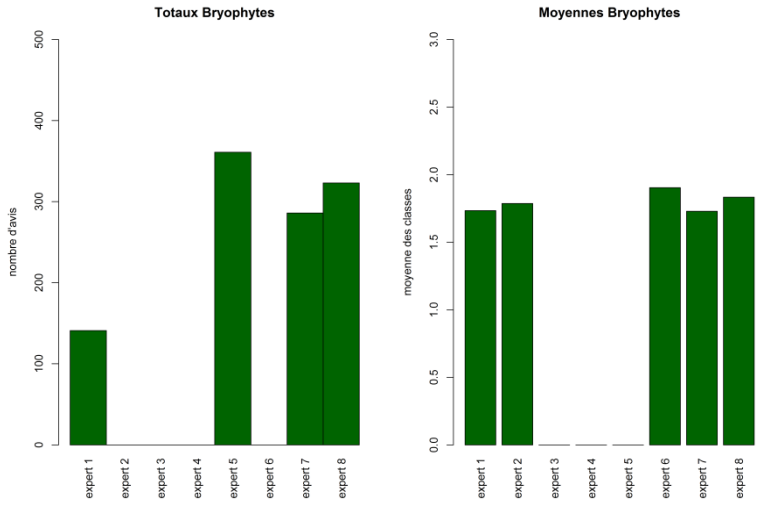


Figure 14 : Différences dans les avis d'experts quant à la quantité de couples traités et les classes attribuées chez les Bryophytes

Analyse de la distribution des confusions

Une fois le maximum parmi les avis ayant été choisi comme coefficient de confusion, le nombre de taxons par nombre total de confusions a également été établi (cf figure 15). Ceci a permis de voir qu'il y avait dans la globalité une cohérence avec une valeur maximale obtenue pour un total de 1 confusion (les taxons avec 0 confusion ont été enlevés du graphique car ils auraient « pollué » les données de par leur occurrence élevée), et le nombre de taxons qui diminue quand le nombre de confusions augmente. Cependant on peut remarquer qu'il y a certaines incohérences, notamment au niveau des confusions au nombre de 4, 6, et 9. Les experts de l'équipe ont alors examiné les taxons concernés et corrigé leurs classes de confusion (qu'ils avaient eux-mêmes attribuées) quand celles-ci s'avéraient mal estimées.

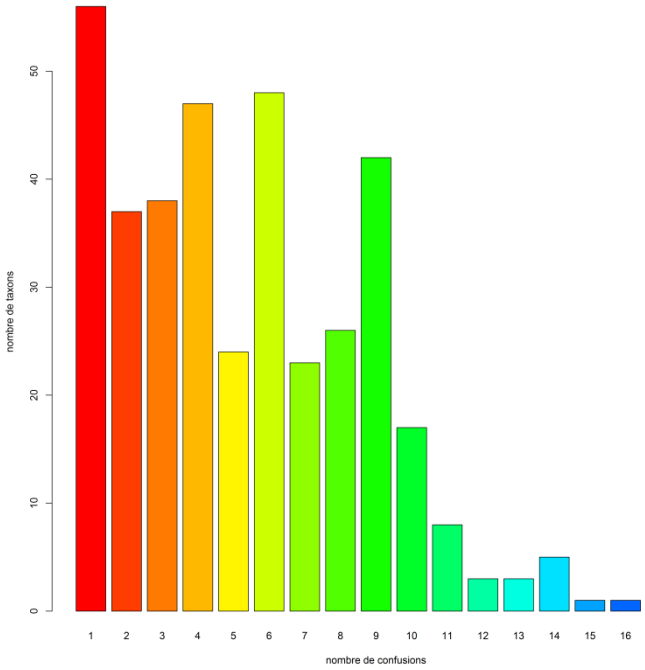


Figure 15 : Distribution des taxons en fonction du nombre de taxons avec lesquels ils sont confondus

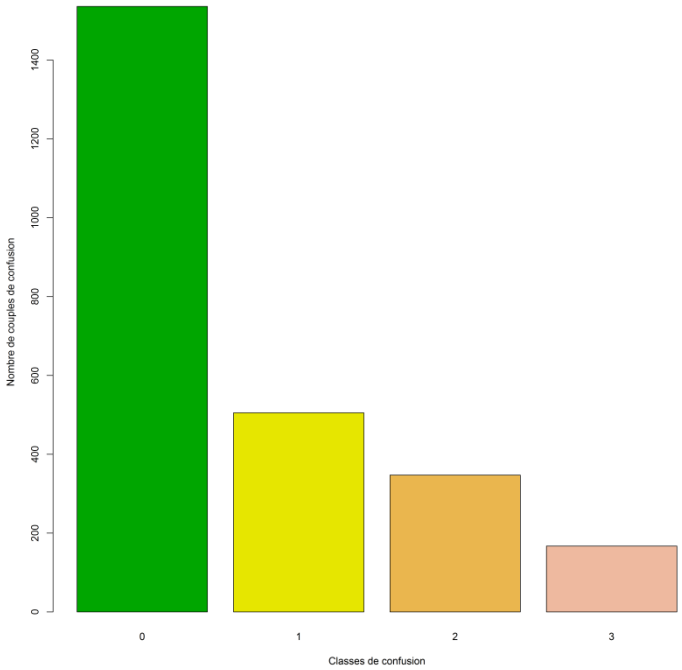


Figure 16 : Distribution des couples de confusion par classe de confusion

Une analyse du nombre de couples taxinomiques par classe de confusion a par la suite été réalisée (cf figure 16), afin de vérifier si la distribution était cohérente. Cette distribution diminuant nettement avec la classe de confusion qui augmente, montre que dans l'ensemble, les classes de confusions ont été attribuées de façon statistiquement logique.

Le tableau 8 montre un aperçu de l'analyse du nombre de taxons, par classe de confusion, avec lesquels chaque taxon peut être potentiellement confondu. Par exemple on peut voir que ACOCAL (*Acorus calamus*) peut être confondu avec 5 taxons par une confusion de classe 1, 2 taxons par une confusion de classe 2, et 0 taxon par une confusion de classe 3.
Ce tableau a permis de déterminer un scénario permettant de recoder les classes de confusion en pourcentages afin de les utiliser dans la matrice de confusion.

Tableau 8 : Nombre de confusions par classe de confusion pour chaque taxon

cd_taxon	classe1	classe2	classe3
ACOCAL	5	2	0
AGRCAN	2	0	2
AGRCAP	1	2	0
AGRCUR	1	0	0
AGRSTO	1	5	2
ALIGRA	2	3	0

Choix du scénario de dégradation

A partir de ce tableau ont pu être testés 3 scénarii possibles, attribuant pour chaque classe de confusion un pourcentage propre afin de traduire ces classes en pourcentages (cf tableau 9).

Tableau 9 : Caractéristiques des 3 scénarii de recodage des classes de confusion

	Classe 1	Classe 2	Classe 3
Scénario 0	0.1 %	1 %	10 %
Scénario 1	1 %	5 %	10 %
Scénario 2	1 %	10 %	25 %

Il a fallu multiplier, pour chaque classe, le nombre de taxons par le pourcentage attribué selon le scénario choisi, et sommer le résultat de chaque classe. Par exemple, on peut voir dans la figure 17 ci-dessous, dans le cadre bleu, que pour ACOCAL il a fallu faire :

Pour le scénario 0 : 5x0.1 + 2x1 + 0x10 = 2.5 % => résultat affiché dans la colonne « pc_conf0 »
Pour le scénario 1 : 5x1 + 2x5 + 0x10 = 15 % => résultat affiché dans la colonne « pc_conf1 »
Pour le scénario 2 : 5x1 + 2x10 + 0x25 = 25 % => résultat affiché dans la colonne « pc_conf2 »
Le résultat pour les mêmes données peut donc extrêmement varier en fonction du scénario choisi.

cd_taxon	classe1	classe2	classe3	motif	pc_conf0	pc_conf1	pc_conf2
ACOCAL	5	2	0	5_2_0	2,5	15	25
AGRCAN	2	0	2	2_0_2	20,2	22	52
AGRCAP	1	2	0	1_2_0	2,1	11	21
AGRCUR	1	0	0	1_0_0	0,1	1	1
AGRSTO	1	5	2	1_5_2	25,1	46	101
ALIGRA	2	3	0	2_3_0	3,2	17	32

Figure 17 : Choix du scénario à utiliser pour recoder les classes de confusion en % pour la matrice de confusion

En s'appuyant sur les valeurs numériques obtenues dans ce tableau, ainsi que sur les résultats illustrés dans le graphique qui en est tiré (cf figure 18), le scénario 1 a pu être choisi car considéré comme plus adéquat.

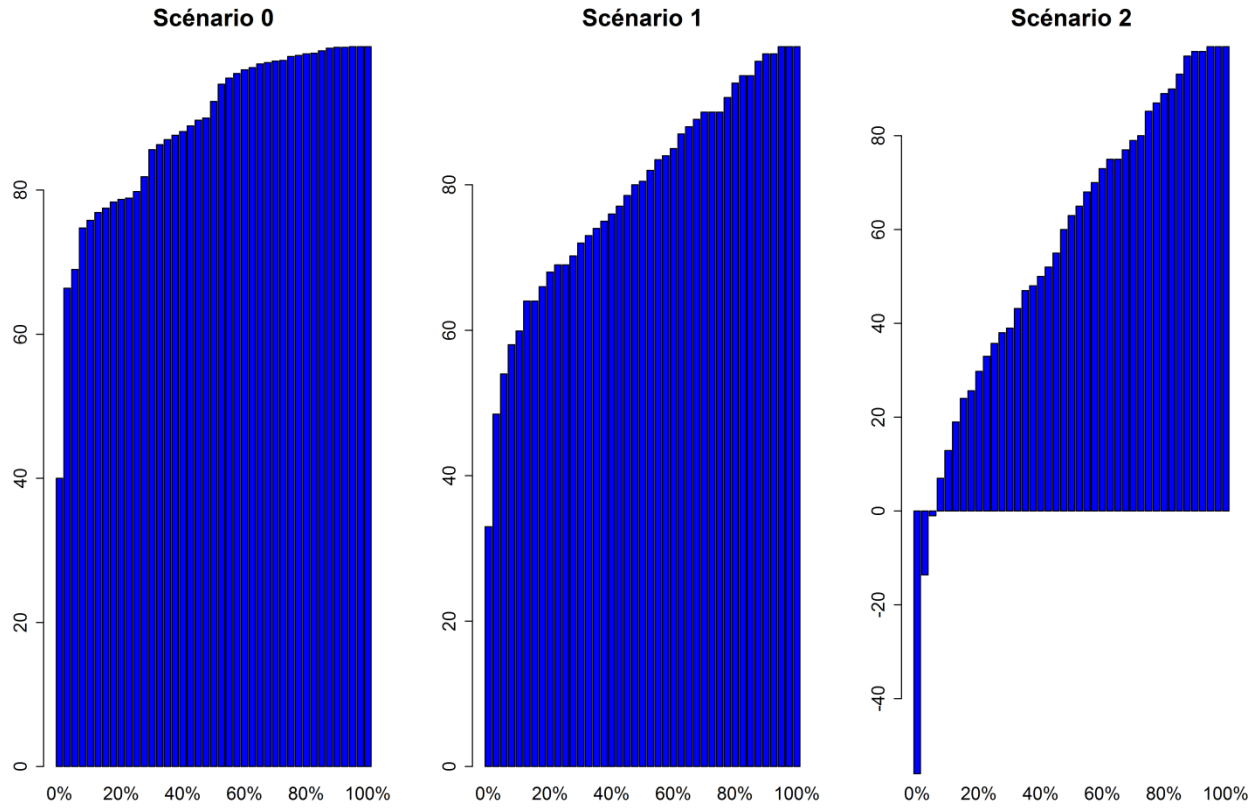


Figure 18 : Distribution des bonnes déterminations pour chaque scénario

En effet, le pourcentage total de confusion ne doit jamais dépasser les 100% car cela constitue un non-sens. De même, ce pourcentage ne doit pas être trop élevé car il ne faut pas oublier qu'il y a généralement plusieurs confusions possibles (à sommer), et surtout qu'il faut prendre en compte qu'il y a aussi une forte probabilité qu'il n'y ait pas confusion (bonne détermination).

Finalement il faut que le pourcentage de confusion soit assez élevé pour qu'il y ait une possibilité de voir certaines confusions apparaître par rapport au nombre de matrices dégradées produites. Si la matrice de confusion (contenant les pourcentages de confusion) n'était pas assez dégradante, les matrices dégradées produites seront toutes similaires à la matrice originale et les confusions qui interviennent dans la réalité ne seraient alors pas représentées.

De cette manière, le scénario 2 se montre trop dégradant (certains taux de bonne détermination sont négatifs), et le scénario 0 ne l'est pas assez (les ¾ des taxons à plus de 80% de bonne détermination). C'est donc le scénario 1 qui a été choisi.

Calibration du nombre de dégradation

Pour s'assurer d'effectuer un nombre de mélange de confusions assez important pour être représentatif de l'impact global de cette matrice, le pourcentage de confusions a été déterminé selon le nombre de dégradation fixé (cf figure 19) a été évalué.

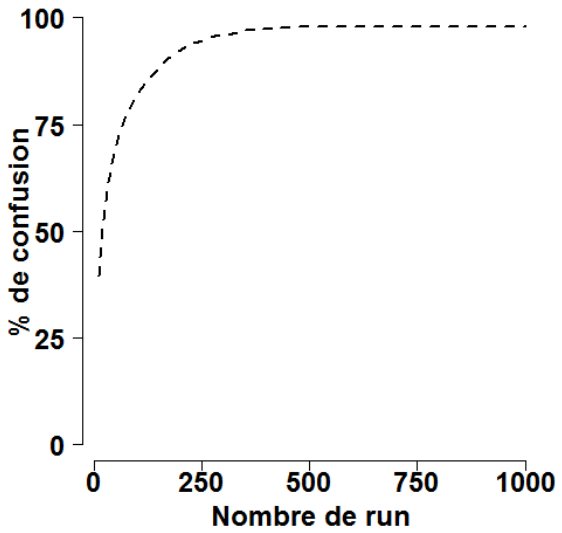


Figure 19 : Taux de confusion pris en compte selon le nombre de dégradation fixé

Il apparaît qu'il est nécessaire d'effectuer au moins 740 dégradations aléatoires pour s'assurer de prendre en compte au moins une fois chacune des confusions présentes dans la matrice de confusion.
Pour assurer cette représentativité, le nombre d'itérations retenu a été de 1,33 fois le nombre d'itérations minimal (+33%), soit 1000 itérations.

Analyse de la distribution des confusions en fonction de la contributivité

Dans les figures 20 et 21, un « motif de confusion» a été associé à tous les couples en fonction de la cote spécifique CS (et donc selon la contributivité) de chacun des taxons, afin de pouvoir faire des graphiques permettant de voir quels types de confusion sont les plus courants. La figure 20 présente ces motifs pour les couples composés de 2 taxons contributifs (avec un CS attribué), et la figure 21 pour les couples avec au moins un taxon non contributif (pas de CS attribué). Ce CS est important car c'est sa présence/son absence qui va impacter le calcul de l'IBMR dans lequel il est utilisé (cf I.1.c. « Les macrophytes en tant que bioindicateurs »).

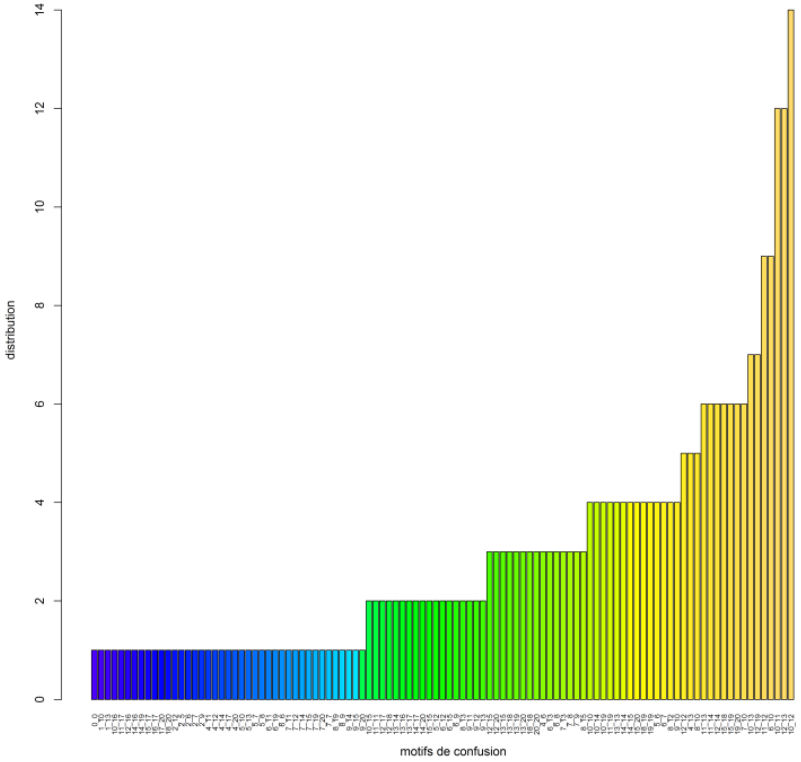


Figure 20 : Distribution des motifs CS des couples de taxons contributifs pouvant être confondus

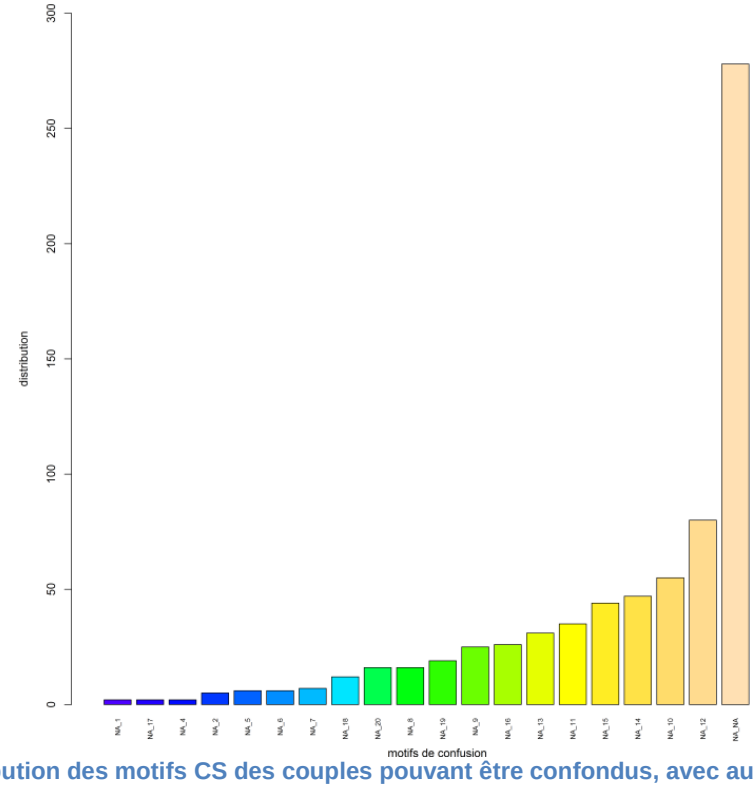


Figure 21 : Distribution des motifs CS des couples pouvant être confondus, avec au moins un taxon non contributif

III- Résultats et perspectives

1) Répertoire de l'analyse bibliographique

L'analyse bibliographique a permis de constituer un répertoire inédit de synthèse des confusions mentionnées dans les différents ouvrages décrivant la flore aquatique [31-39].

Ce répertoire a été produit sur Excel sous 2 formes :

- 1 forme rassemblant en plus des couples et remarques, les sources bibliographiques précisées pour chaque couple de confusion taxinomique

Tableau 10 : Aperçu de la forme avec sources bibliographiques du répertoire des couples de confusion

Confusion possible			Sources bibliographiques		
Taxon A	Taxon B	Remarques	(Bailly et al., 2010)	(Mouronval et al., 2010))	(Spencer-Jones et al., 1986)
Chara braunii Gmel.	Nitellopsis	En ce qui concerne les petits Nitellopsis.	X		

- 1 forme permettant l'analyse informatique des couples, rassemblant en plus des couples et remarques, les codes 6L de chaque taxon de chaque couple de confusion taxinomique

Tableau 11: Aperçu de la forme avec codes 6L du répertoire des couples de confusion

Taxon A	Taxon B	Remarques	code 6L taxon A	code 6L taxon B
Chara braunii	Nitellopsis	En ce qui concerne les petits Nitellopsis.	CHABRA	NIESPX

Ainsi 758 taxons ont été mis en évidence pour les 5 groupes taxinomiques présentés dans le tableau 12 ci-dessous. Les groupes les plus représentés sont les Bryophytes, puis les Phanérogames, suivis des Ptéridophytes.
Seuls les groupes taxinomiques Hétérotrophes et Lichens n'ont pas été représentés puisqu'ils sont très peu présents dans la base de données Pandore. De plus, la bibliographie en rapport avec ces taxons étant très dense, il a semblé inutile de consacrer un temps trop long dans cette étude bibliographique pour les quelques taxons présents dans Pandore.
Entre 2 et 3 guides de détermination ont été utilisés pour chaque groupe taxinomique, pour un total de 9 ouvrages pour l'ensemble du répertoire, certains ouvrages rassemblant différents groupes taxinomiques.

De façon chiffrée, le répertoire correspond au tableau suivant :

Tableau 12 : Le répertoire des couples de confusion en chiffres

Groupe taxinomique	Nombre de couples de confusion	Pourcentage sur le total des couples (%)	Nombre d'ouvrages utilisés
Algues	97	12.8	2
Bryophytes	227	29.9	2
Characées	49	6.5	3
Phanérogames	196	25.9	3
Ptéridophytes	189	24.9	3
Total	758	100	9

2) Description de la matrice de confusion taxinomique

Suite à l'enquête auprès des experts, la matrice de confusion taxinomique a été conçue dans un format utilisable pour les analyses numériques. La figure ci-dessous illustre une partie de cette matrice :

X	ACOCAL	AGRCAN	AGRCAP	AGRCUR	AGRSTO	ALIGRA	ALILAN	ALIPLA	ANASPX	ANEPIN	ARUDON	AUDSPX
ACOCAL	85	0	0	0	0	0	0	0	0	0	0	0
AGRCAN	0	78	0	0	10	0	0	0	0	0	0	0
AGRCAP	0	0	89	0	5	0	0	0	0	0	0	0
AGRCUR	0	0	0	99	1	0	0	0	0	0	0	0
AGRSTO	0	10	5	1	54	0	0	0	0	0	0	0
ALIGRA	0	0	0	0	0	83	5	5	0	0	0	0
ALILAN	0	0	0	0	0	5	74	10	0	0	0	0
ALIPLA	0	0	0	0	0	5	10	77	0	0	0	0
ANASPX	0	0	0	0	0	0	0	0	64	0	0	0
ANEPIN	0	0	0	0	0	0	0	0	0	59	0	0
ARUDON	0	0	0	0	0	0	0	0	0	0	95	0
AUDSPX	0	0	0	0	0	0	0	0	0	0	0	94

Figure 22 : Extrait de la matrice de confusion taxinomique

Cette matrice est donc composée des coefficients de confusion entre les différents taxons. La diagonale représente la « confusion » des taxons avec eux-mêmes, c'est-à-dire la probabilité de « bonne détermination ». La somme de chaque ligne fait 100%, il en va de même pour la somme de chaque colonne. Cette matrice est composée de 460 taxons, et donc de 460 x 460 = 211.600 coefficients de confusion.

On peut voir dans la figure 23 et le tableau 13 que le groupe des Phanérogames est celui contenant le plus de confusions possibles, suivi par les Algues (dont Characées) et les Bryophytes, ce qui correspond globalement aux occurrences de ces groupes taxinomiques observées dans la base de données Pandore. Les pourcentages représentent le taux de couples de confusion parmi l'ensemble des couples de confusion de tous les groupes taxinomiques réunis.

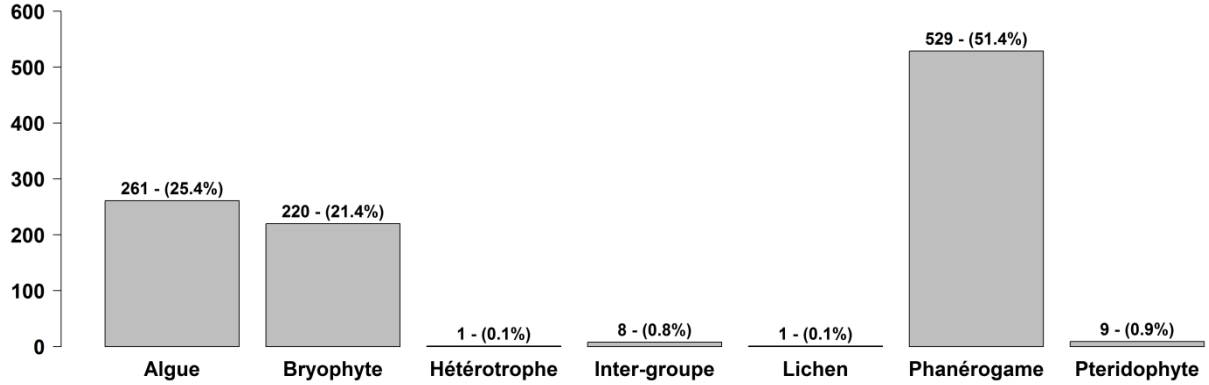


Figure 23: Nombre de couples de confusion par groupe taxinomique

Le nombre de taxons pouvant être confondus a aussi été consigné dans le tableau 13. On peut constater qu'il y a au total 460 taxons pouvant être sujets à confusion, et que pour la majorité (55.9%) il s'agit des phanérogames.

Tableau 13: Nombre de couples de confusion et nombre de taxons pour chaque groupe taxinomique

Groupe taxinomique	Nombre de couples de confusion	Pourcentage sur le nombre total de couples (%)	Nombre de taxons dans le groupe taxinomique	Pourcentage sur le nombre total de taxons (%)
Algues (dont Characées)	261	25.3	109	23.7
Bryophytes	220	21.4	80	17.4
Hétérotrophes	1	0.1	2	0.4
Lichens	1	0.1	3	0.6
Phanérogames	529	51.4	257	55.9
Ptéridophytes	9	0.9	9	2.0
Inter-groupes	8	0.8	/	/
Total	1029	100	460	100

Il y a un enjeu particulier pour les taxons contributeurs au calcul de l'IBMR, c'est-à-dire ceux ayant une côte spécifique CS (cf *I.1.b. « Les bioindicateurs, des outils d'évaluation de l'état écologique »*).
Dans cette matrice de confusion, sont inclus dans au moins une confusion, 170 taxons sur les 208 contributeurs au calcul de l'IBMR, soit 81%, ce qui laisse présager un impact potentiel important sur l'IBMR de cette matrice.

Les taxons indicateurs utilisés dans la matrice de confusion sont représentés dans les Annexe 4 (tableau) et 5 (graphique). Ils ont été triés en fonction de leur taux de bonne détermination. Ainsi on peut voir que le taxon le moins soumis à confusion (taux de bonne détermination le plus élevé) serait *Vaucheria sp.* (Algue) avec 99% de bonne détermination, suivi de *Tetraspora sp.* (Algue) et de *Riccia fluitans* (Bryophyte), tous deux à 98% de bonne détermination. En revanche le taxon le plus soumis à confusion est *Nitella flexilis* (Characée) avec 33% de bonne détermination, suivi de *Nitella gracilis* (Characée) avec 34% de bonne détermination, puis de *Phormidium sp.* (Algue) avec 36% de bonne détermination.

A partir de cette matrice et des listes floristiques provenant de Pandore, 2.629.000 listes dégradées ont été créées, comprenant l'ensemble de ces 1029 confusions taxinomiques.

3) Impact des confusions par approche unitaire

L'impact sur l'IBMR de chacune des 1029 confusions taxinomiques a été étudié sur l'ensemble des listes floristiques du jeu de données.

La figure 24 ci-dessous représente les couples ayant l'impact le plus fort sur l'IBMR (écart <-1 en rouge ou >1 en bleu). En plus de cet impact, il y a d'autres critères importants indiqués pour chaque taxon analysé (qui est le premier écrit dans le couple) : le nombre d'occurrence du taxon (n=…), ainsi que la probabilité de la confusion (…%).

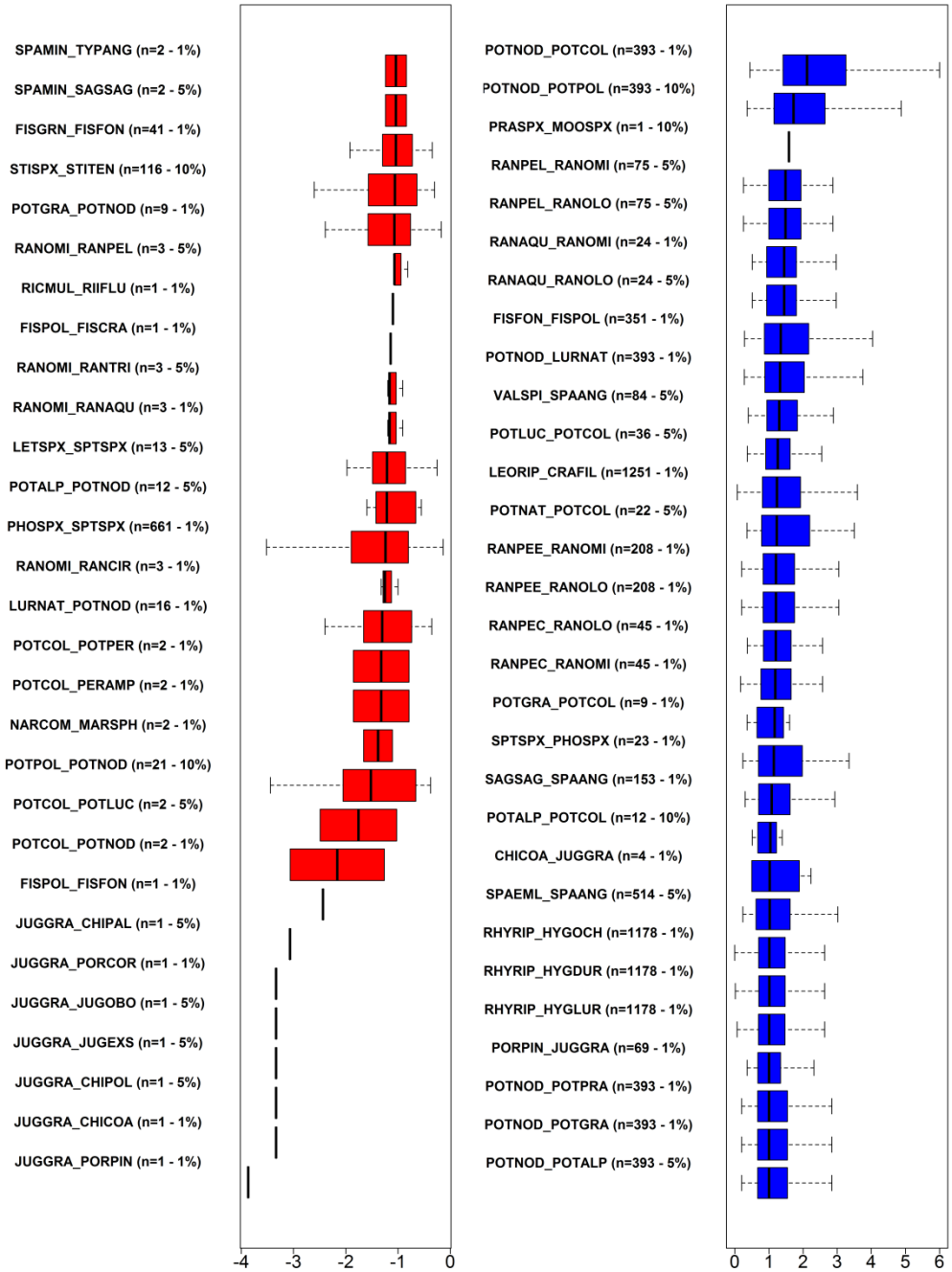


Figure 24 : Impact des couples de confusion altérant significativement l'IBMR de façon négative (en rouge à gauche) et positive (en bleu à droite)

Il apparaît que seules 29 confusions ont un impact médian altérant négativement l'IBMR de plus de 1 unité et 30 ont un impact positif de plus de 1 unité.

Ainsi on peut voir que le couple ayant le plus fort impact « positif » (c'est-à-dire écart d'IBMR>0) est POTNOD_POTCOL. Le taxon analysé est donc POTNOD (*Potamogeton nodosus*), et on peut voir que l'impact de cette confusion avec POTCOL (*Potamogeton coloratus*) sur l'IBMR est très fort (>+2) et que POTNOD apparait souvent dans les listes floristiques, ce sont donc des caractéristiques laissant à penser qu'il faut faire très attention à cette confusion lors des relevés IBMR. Cependant on peut voir que le risque de faire cette confusion est de 1%, ce qui est très faible et donc a peu de probabilités de se produire ; ce n'est donc pas le couple auquel il faut le plus prêter attention.
En revanche, le couple POTNOD et POTPOL (*Potamogeton polygonifolius*) qui présente un écart d'IBMR>1.5, une occurrence de 393 pour POTNOD et une probabilité de confusion de 10%, a plus de risque d'avoir lieu tout en ayant un fort impact sur l'IBMR. Les opérateurs devront donc faire très attention à cette possibilité de confusion.

Il en va de même pour les écarts d'IBMR négatifs : plus la valeur absolue de l'écart d'IBMR est élevée, plus l'impact est fort. Le nombre d'occurrence du taxon renseigne sur l'impact de la confusion à l'échelle de l'ensemble des listes floristiques. Une confusion d'un taxon rare, même si sa probabilité est forte et son impact sur l'IBMR important aura un impact global faible du fait de sa rareté. Au contraire un taxon fréquemment rencontré, s'il peut être confondu, aura un impact global d'autant plus important qu'il est occurrent. Ainsi des couples tels que *Jungermannia gracillima* et *Porella pinnata* (JUGGRA_PORPIN) ayant un fort impact sur l'IBMR (presque égal à -4), ne va pas vraiment nécessiter de soin particulier car il n'apparait qu'une fois dans les listes floristiques, et n'a par ailleurs que 1% de risque d'occurrence.
De même des couples ayant un nombre élevé d'occurrence et de probabilité de confusion, mais n'ayant pas d'impact sur l'IBMR (≈0) vont moins nécessiter d'attention.

Il faut donc se baser sur l'ensemble de ces 3 critères (occurrence du taxon, probabilité de confusion, impact sur l'indicateur) pour déterminer les couples les plus à risque pour l'IBMR.

4) Impact des confusions par approche multiple

a. Etude d'impact global sur l'IBMR

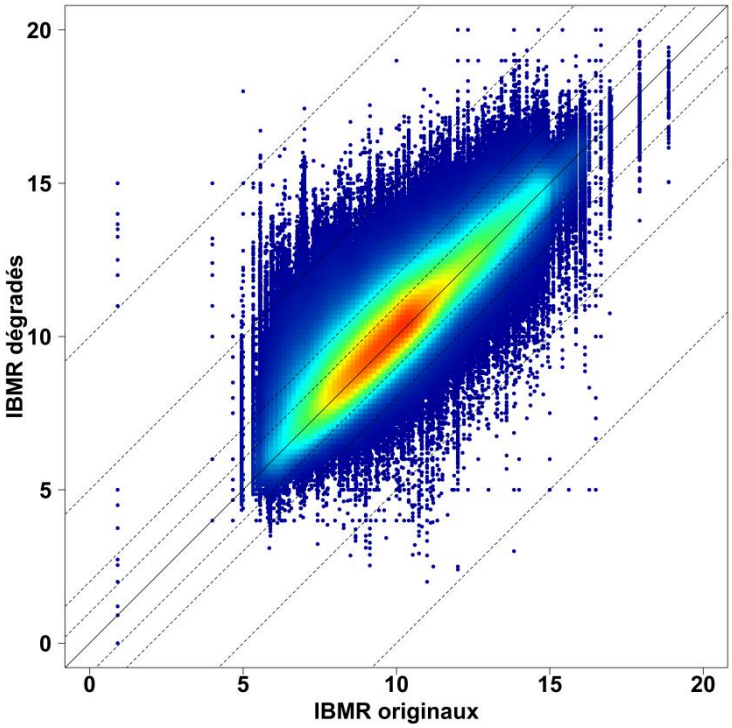


Figure 25 : IBMR originaux en fonction des IBMR dégradés et couleur selon densité de point (rouge : forte densité, bleu : faible densité).

En comparant les IBMR calculés sur les listes floristiques originales à ceux calculés sur les listes dégradées (figure 25), il apparaît que la densité de points la plus forte se situe le long de la droite x=y correspondant aux IBMR non impactés.

Sur la carte de France ci-contre (Figure 26), ont été placées les différentes stations du réseau étudiées. A chacune a été attribuée une couleur en fonction de l'impact maximum sur l'IBMR obtenu sur l'ensemble des dégradations faites sur ce site.

On peut voir que dans la majorité des cas il s'agit d'un écart d'IBMR compris entre 0 et 4,36 unités (bleu et vert).
Grâce à ce type d'illustration des résultats obtenus par la dégradation des listes floristiques, il est possible d'estimer des zones et sites à plus ou moins de risques de confusions taxinomiques. La partie Sud-Ouest de la France, aux alentours de la Garonne, semble par exemple plus sujette à des impacts élevés sur l'IBMR (beaucoup de rouge, jaune, et orange), alors que la partie Nord-Ouest, englobant la Bretagne et ses environs, semble au contraire plus encline aux écarts d'IBMR faibles (surtout du bleu et du vert).

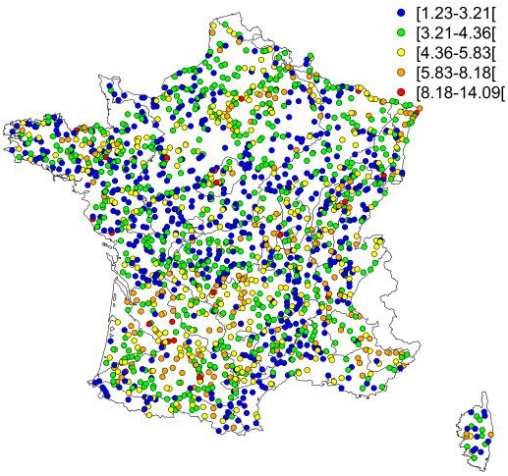


Figure 26 : Cartographie de la valeur absolue de l'impact maximum sur l'IBMR de l'ensemble des dégradations (en rouge : impacts les plus forts, en bleu : les moins forts)

L'observation de la médiane de l'impact sur l'IBMR des dégradations par site (cf figure 27) montre qu'une majorité de site (68%) ont plus de 50% de leur dégradation impliquant une augmentation supérieure à 0.1 unité IBMR, et très peu (3%) un impact négatif (inférieur ou égal à -0.1 unité IBMR). Ces résultats montrent donc que la confusion taxinomique, selon l'approche qui a été menée, entraîne une tendance à une élévation globale des valeurs d'IBMR.

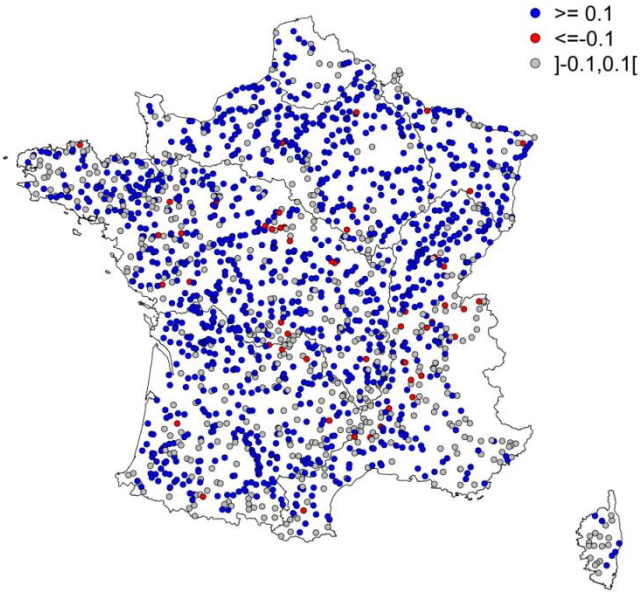


Figure 27: Impact médian de l'ensemble des confusions sur l'IBMR

Ce résultat global de tendance à l'augmentation des valeurs de l'indicateur n'a pas d'explication *a priori*.
Une des hypothèses principales expliquant ce phénomène est le choix méthodologique qui a été fait dans cette première approche de considérer les confusions comme étant symétriques. En effet, en remplaçant de façon équivalente des taxons très fréquents et à cote spécifique CS moyenne ou faible par des taxons confondables mais rares et à CS élevé, on introduit un biais en surévaluant un des deux sens de l'impact potentiel du couple de confusion. Lorsque l'on examine l'impact unitaire des couples , on voit que les couples qui ont le plus fort impact sont en effet très dissymétriques à la fois dans leur occurrence et dans leur valence écologique.

Ce modèle d'impact dans l'indicateur devra donc être affiné par un système de pondération traduisant la probabilité du sens de confusion, afin de corriger ce biais.

b. Impact en termes d'évaluation de l'état écologique

L'objectif opérationnel ultime dans le contexte méthodologique DCE étant l'évaluation de l'impact sur le classement d'état écologique, il est intéressant de voir si cet impact est le même ou diffère sur le classement.
La création d'un tableau de contingence permet d'estimer cet impact, à partir du jeu de données d'IBMR altérés.

Dans le tableau 14 sont indiqués les différents taux de classements différents ou identiques pour l'ensemble des données répartis par classes de qualité. La diagonale représente le taux de bon classement (isoclassement) de chacune d'elles, c'est-à-dire de classement dans la même classe d'état que celle donnée par la liste floristique originale. Au-dessus de cette diagonale se trouvent les sur-classements (la classe d'état attribuée est plus haute que la classe originale), et en-dessous les déclassements (la classe d'état attribuée est plus basse que la classe originale).
On peut voir que, pour la majeure partie des relevés (68.35% des cas), la classe d'état obtenue à partir d'une liste altérée est la même que celle indiquée avec la liste originale. C'est donc 31,65% des classements qui sont modifiés par une ou des confusions introduites dans les listes. On remarque à l'aide de ce tableau et de la figure 28 que la probabilité de confusion est très majoritairement entre classes voisines (28.3% pour un sur/déclassement de +/1 classe, 2.9% pour +/-2 classes, 0.3% pour +/- 3 classes, 0% pour +/- 4 classes).

Tableau 14 : Taux de classements identiques et différents

		Classement original				
		Très bon	Bon	Moyen	Médiocre	Mauvais
Classement dégradé	Très bon	27.58	11.42	1.74	0.28	0.04
	Bon	2.75	24.23	7.5	0.74	0.05
	Moyen	0.13	2.59	12.14	2.33	0.11
	Médiocre	0.02	0.13	1.19	3.67	0.25
	Mauvais	0.01	0.03	0.05	0.23	0.73

D'après le tableau 14 et la figure 28, on peut constater que le taux de surclassement (24.46%) est largement plus élevé que le taux de déclassement (7.13%), ce qui ne correspond pas aux observations de J. Wiederkehr. [27]

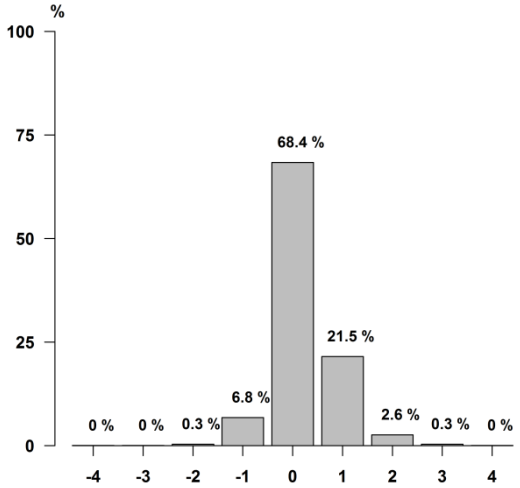


Figure 28: Distribution du nombre de changement de classe sur l'ensemble des confusions

La limite de classe entre Bon et Moyen est la limite la plus importante au sens de la DCE de par les exigences faites au niveau de la qualité des eaux.
Dans le graphique suivant, c'est donc les classements en fonction de cette limite qui ont été analysés.

c. Régionalisation de l'impact sur le classement

Les différences de classements entre régions écologiques sont illustrées par la répartition de ces classements en fonction des biotypes.
Les biotypes correspondent à des types de peuplements macrophytiques caractéristiques des différents types de cours d'eau rencontrés en France. Ces biotypes, au nombre de 8, ont été définis par l'Irstea par analyse de l'ensemble des relevés de végétation aquatique sur un réseau de sites de référence, dans le cadre de l'élaboration de la méthode nationale d'évaluation. [18]

Dans chacun des cas, il y a au moins 80% de bonne classification (quelle que soit la classe), et une majorité de surclassement (passage de classe inférieure à la limite de Bon état (BE) à une classe supérieure à la limite de Bon état) par rapport au déclassement (passage de classe supérieure à la limite de Bon état à une classe inférieure à la limite de Bon état), ce qui correspond aux données du tableau précédent.

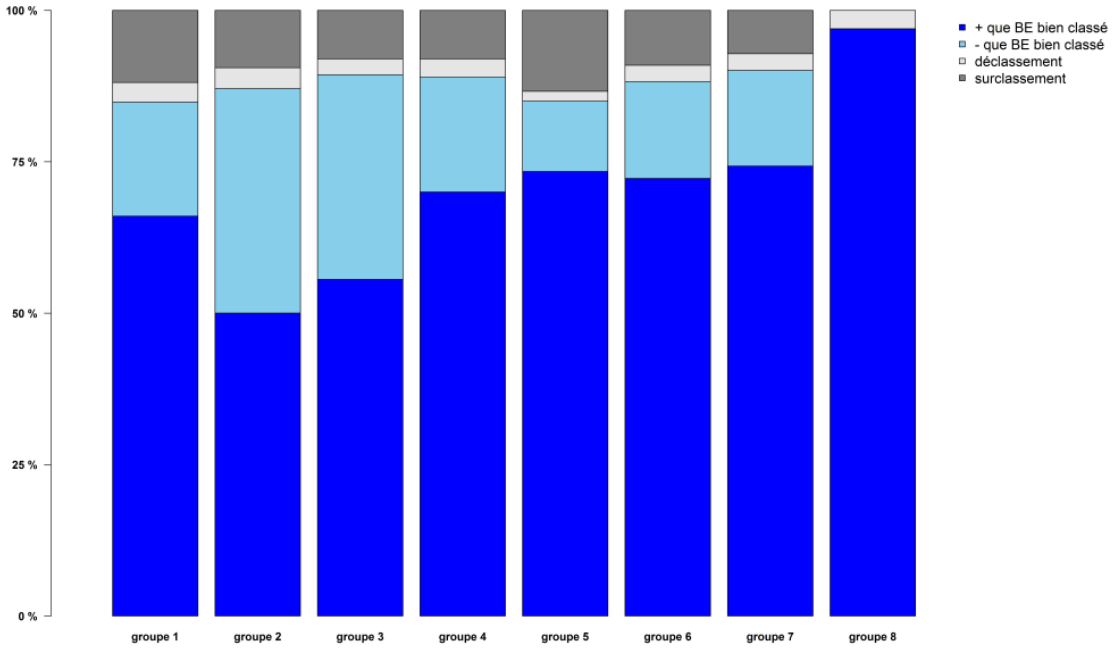


Figure 29 : Proportions de bonne et mauvaise classification en fonction de la limite de Bon état/ état Moyen

Les taux de bons classements sont globalement comparables d'un biotype à l'autre (80%), avec de légères différences. Le groupe 5 (cours d'eau grands à moyen mésotrophes à eutrophes), par exemple, montre un taux de bon classement un peu inférieur, très principalement dû aux classes inférieur au Bon état.

En revanche certaines différences subsistent au sein de ces bons classements : la composition en classes supérieures/inférieures à la limite Bon état diffère, ce qui est très probablement lié à la différence dans la composition taxinomique, ainsi qu'à l'état écologique général de ces types de cours d'eau, qui n'est pas équivalent selon ces types.
Le biotype 8 (très grands cours d'eau) montre des caractéristiques différant des autres par l'absence de classes inférieures à la limite de Bon Etat, d'où l'absence de surclassements et de classes inférieures à Bon état bien classées. Ceci est lié aux taxons caractéristiques de ce biotype.

5) Evaluation de l'étude

a. Discussion

L'approche unitaire apporte un intérêt pour identifier les couples présentant un risque de confusion qui va impacter les classes d'état des cours d'eau. Ainsi, les couples les plus à risques sont ceux présentant un fort impact sur l'IBMR (positif ou négatif), mais aussi ayant une forte occurrence dans les listes floristiques (et donc dans les différents sites étudiés), et un coefficient de confusion élevé.
Il faut donc se baser sur l'ensemble de ces 3 critères pour déterminer les couples les plus à risque pour l'IBMR.

Pour comparer les impacts sur l'IBMR et ceux sur les classes d'état, il faut garder à l'esprit la nature différente de réponse potentielle, ainsi que les enjeux entre :

- Indicateur (l'IBMR): il s'agit de valeurs continues, et la vision du niveau trophique est potentiellement biaisée par les confusions,
- Etat écologique : il s'agit de classes définies par des limites précises en fonction des valeurs d'EQR (cf partie I.1.c c. « Les macrophytes en tant que bioindicateurs »), et la définition de ces classes présente un enjeu de gestion.

L'impact des confusions taxinomiques sur l'IBMR a montré une tendance majoritaire à l'augmentation, plutôt qu'à la diminution ou qu'à l'obtention de valeur similaire (entre -0.1 et +0.1 unités).
Cette tendance se retrouve logiquement dans l'impact sur le classement d'état écologique résultant. Elle constitue très probablement un artéfact lié au choix fait pour cette première approche de ne pas considérer l'asymétrie des confusions taxinomiques.

La comparaison avec les travaux antérieurs focalisés sur l'impact des confusions taxinomiques (Wiederkehr 2015 .[27, 28]) montre des résultats assez différents. Plusieurs raisons peuvent être avancées pour expliquer ces différences :

- Le jeu de données utilisé dans les travaux de Wiederkehr était issu de mesures sur une zone géographique plus limitée (France Nord- Ouest), d'où des types de peuplements végétaux et des probabilités de confusion sans doute un peu différents. De plus, comme on a peut le voir par l'approche régionalisée (cf figure 26), la région concernée (la partie Nord-Ouest de la France) correspond à une zone où les écarts d'IBMR paraissent plus faibles par rapport à l'ensemble des sites observés.
- Le nombre de taxons considéré est plus grand dans notre étude (1029 dont 460 pouvant être confondus) que dans les travaux de J. Wiederkehr (219 dont 82 pouvant être confondus), ce qui peut entraîner une différence dans les résultats. Cette différence de taxons utilisés est aussi lié au choix de la zone d'étude plus restreinte.
- La non prise en compte des asymétries de confusion dans notre étude a aussi joué un rôle certain dans l'amplification du risque de confusion vers un taxon à CS plus élevé. Afin d'affiner le modèle, il faudra alors envisager la prise en compte de cette asymétrie.

Finalement, l'impact des confusions peut être significativement différent en fonction du biotype, et donc du type de cours d'eau ou de l'HER (hydroécorégions, au sens de Wasson et al. 2006 [42]). Rappelons que ces HER constituent la base de la définition des types biologiques et des valeurs de référence pour la plupart des indicateurs utilisés dans l'évaluation des masses d'eau. Cette régionalisation constitue une piste de réflexion dans l'optique d'affiner l'utilisation des résultats de l'approche « confusions ».

b. Bénéfices environnementaux, économiques et sociétaux des opérations

Le bénéfice sociétal de ce travail rejoint le bénéfice environnemental : la bonne qualité des eaux.
En effet, par ce travail sur les incertitudes, les décisions pourront être prises pour l'application d'actions sur les différents cours d'eau nécessitant réellement une amélioration de leur qualité.

Ces mesures faites dans le but d'assurer un bon état des eaux au niveau national ont un coût économique important. Cette dépense, correspondant au coût global des programmes de mesures, est en effet estimée à 27 milliards d'euros pour la période 2009-2015 [43]. La mobilisation de tels budgets nécessite un niveau de confiance en conséquence, confiance qui dépend, entre autres, de la précision et de la fiabilité des méthodes d'évaluation, pour l'état et le diagnostic. En effet, si un cours d'eau est par exemple classé juste en-dessous de la limite Bon Etat avec 40% de confiance, la pertinence des investissements peut être discutée.
C'est pourquoi une connaissance avancée sur les incertitudes est primordiale, afin que le budget puisse être ciblé sur les cas qui le nécessitent réellement, avec le moins d'erreur possible.
Une confiance élevée dans les mesures optimise ainsi la dépense d'argent.

De même, l'évaluation de l'état écologique des masses d'eau revêt un enjeu politique. En allant plus loin que la prescription européenne qui fixe à 50% de bon état en 2015, la loi « Grenelle II » [44] fixe un objectif de 66% de bon état en 2015. La fiabilité dans l'évaluation impacte inévitablement le résultat obtenu, du moins dans son résultat affiché, et donc l'atteinte ou non des objectifs fixés par la loi pour chaque bassin.

Le travail sur les incertitudes s'inscrit dans ce cadre : contribuer à identifier les sources d'incertitude des méthodes d'évaluation et estimer leur impact, afin de proposer des protocoles de prise en compte de ces niveaux de confiance ou des ajustements correctifs dans les méthodes.

c. Analyse des écarts des résultats

La préparation du document de consultation, à envoyer aux experts, a pris plus de temps que ce qui était prévu dans le planning en raison du prétraitement à réaliser. Ce retard s'est allongé avec le temps de retour des fichiers remplis qui a, lui aussi, été plus long que prévu dans le planning.
Ceci illustre un problème pouvant être rencontré lors de travail avec plusieurs intervenants : les délais deviennent plus difficiles à maîtriser. De plus, malgré l'harmonisation faite au niveau du fichier et des instructions, et avec le besoin de faire un travail complet et consensuel, de nouveaux problèmes sont apparus, rendant plus difficile la réalisation des objectifs dans les temps impartis.

Le manque de temps a aussi impliqué que le module complet fini ne pouvait être totalement atteint au bout des 6 mois de stage. Cependant le planning était conçu de manière à gérer l'éventualité d'une non coopération des experts, ce qui aurait entraîné une impossibilité de réaliser la matrice de confusion. Ainsi, dans de telles conditions, la mission de stage aurait plus été tournée vers l'étude de la « variabilité stationnelle ».

d. Devenir des résultats

Ce travail consiste en un travail préparatoire et exploratoire par rapport au programme d'évaluation des incertitudes liée à la confusion taxinomique.
Il sera possible d'intégrer de nouvelles données pour obtenir de nouveaux résultats plus consensuels et complets, tant sur le nombre d'experts participant que sur la quantité de taxons étudiés, ou sur la complexité de la méthode utilisée.
La matrice de confusion a vocation à être partagée avec la communauté scientifique pour poursuivre l'analyse des impacts. De ce travail, des bénéfices dérivés pourront aussi être tirés : formation des opérateurs, identification des groupes taxinomiques pour lesquels il faudra plus de documents de détermination, adaptation du contenu des audits « qualité », etc.
Ces travaux seront poursuivis par l'équipe Irstea de Bordeaux, sur la base de ces premiers résultats et des points à développer qui ont été identifiés pour mettre au point un modèle réaliste et précis d'évaluation des impacts des confusions.

Conclusion

En fonction de la région, du biotype, les écarts d'IBMR observés par rapport aux valeurs originales vont varier d'un site à l'autre. On obtient donc tout un gradient de variations des indicateurs d'état écologique en jeu. Il n'est donc pas possible, ni intéressant, de répondre à la problématique de façon généralisée.
Ainsi, suite à ce travail ont été déterminés les différents écarts obtenus selon les couples de confusion, les biotypes, l'emplacement géographique, etc. De cette façon les taxons et autres caractéristiques (biotype, situation géographique, occurrence, etc.) les plus à risques pour l'IBMR ont été identifiés. L'ensemble des résultats pourra être utilisé afin de mettre en garde les observateurs contre les cas à risque pour l'évaluation de l'état écologique.

Nous avons travaillé sur une métrique (IBMR), mais l'obtention/l'absence d'un impact sur cette métrique ne signifie pas forcément que cet impact va être le même pour une autre métrique.
Il ne faut donc pas se cantonner aux seuls résultats de l'IBMR pour résumer l'impact de la confusion taxinomique. D'autant plus que, plus tard, la métrique nationale sera probablement une autre que l'IBMR.

Afin d'identifier les taxons ayant été confondus, il y a aussi possibilité d'utiliser une approche avec une autre métrique descriptive. Par exemple, il y a possibilité d'évaluer une confusion par l'étude du profil écologique par rapport à un indicateur (par ex : NH_4).
De cette façon, si une espèce ne possède pas le même profil écologique que les autres (qui elles ont un profil écologique similaire), alors elle a probablement été confondue avec une autre (qui elle, a le profil écologique correspondant).

Certaines améliorations pouvant être apportées à la démarche ont été identifiées :

- La matrice est faite de façon à caractériser les confusions de façon symétrique bien que ce ne soit pas toujours le cas dans la réalité. Cette limite a été établie sciemment en raison des limites de temps. En effet, considérer les couples dans les 2 sens, et donc de façon asymétrique, aurait demandé trop de temps de réflexion aux experts et leur aurait rajouté du travail supplémentaire.
  En outre ; il peut parfois y avoir confusion dans un sens mais pas dans l'autre. Par exemple des plantes à longues feuilles jeunes peuvent être confondues avec des plantes à feuilles courtes, mais pas l'inverse.
  Pour pallier à ces complications, nous avons donc cherché à maximiser le risque de confusion parmi les avis pour essayer de conserver l'incertitude.

- La possibilité de confusion peut totalement varier selon la région concernée. En effet certaines espèces ne se trouvent pas dans certaines régions, la confusion devient alors quasiment impossible. Il faudrait donc pouvoir déterminer les coefficients de confusion en fonction des différentes zones de répartition des plantes.

- Un travail supplémentaire qui aurait pu être demandé aux experts aurait été de définir le taux de non confusion (/bonne détermination) pour chaque taxon. Ceci aurait permis de faciliter le codage des classes de confusion en % pour les intégrer dans la matrice, et aurait ainsi plus justifié les pourcentages obtenus.

- Avec l'apport d'avis de plus experts, la matrice aurait été encore plus consensuelle et complète.

- Finalement, le fait d'avoir plus de temps à consacrer à cette matrice, autant pour nous que pour l'ensemble des experts, aurait permis d'apporter toutes ces modifications et ainsi d'obtenir une matrice qui mène à des impacts sur l'évaluation de l'état écologique encore plus proches de la réalité.

En terme de valeur ajoutée, nous pouvons dire que nous avons pu aller plus loin que les travaux antérieurs dans le domaine, qui ont été faits sur l'impact des incertitudes sur les indicateurs, en abordant l'impact des confusions taxinomiques sur les classes d'état et donc l'évaluation de l'état écologique des cours d'eau.

La fiabilité des valeurs obtenues dans le cadre de la surveillance DCE se fait grâce au travail sur les incertitudes. Par cette mission, nous avons pu participer à la récolte des connaissances dans le domaine vaste qu'est celui des incertitudes.

En ce qui concerne les leçons personnelles que j'ai pu tirer de ce stage, on peut tout d'abord parler du monde du travail dans le domaine de la recherche. En effet, j'ai eu l'opportunité de travailler au sein d'une équipe de 4 personnes (5 avec moi) où chacun avait sa spécialité et ses compétences et pouvait ainsi faire bénéficier l'ensemble du groupe de sa plus-value. Ce type d'échanges n'est pas aussi bien représenté dans une école où tout le monde suit la même formation. Il s'agit donc d'un partage enrichissant où chacun apporte sa vision et son avis sur les différentes caractéristiques d'une mission, ce qui permet le débat et l'obtention de résultats plus complets.

De plus il y a eu de nombreuses réunions de l'équipe CARMA au sein de l'établissement qui permettaient de se tenir à jour par rapport à la situation de l'équipe CARMA, de l'unité EABX, mais aussi de l'Irstea dans sa globalité. Au cours de ces réunions nous étions aussi informés de l'avancée dans le travail de chacun, et ainsi dans l'avancée scientifique. Au cours de ces présentations, tout le monde participe, ce qui montre un partage, une collaboration, entre le travail des uns et des autres, mais aussi un apport de suggestions et d'aides afin que chacun puisse avancer dans son travail de la façon la plus optimale.

J'ai également eu l'opportunité de participer à une réunion du groupe de travail « Incertitudes » avec l'ONEMA à Paris où j'ai pu rencontrer différentes équipes travaillant aussi sur la problématique des incertitudes. J'ai ainsi pu prendre connaissance des démarches de chacun selon leur problématique (Indice Ichtyo Lac, incertitudes physico-chimiques, etc.), et de l'avancée des connaissances dans ce domaine.

Grâce à ce stage j'ai pu développer mes connaissances dans le domaine de l'écologie et de la qualité de l'eau avec l'utilisation de l'Indice Biologique Macrophytique en Rivière. J'ai ainsi pu plus apprendre sur les macrophytes et de leur intérêt dans l'évaluation de l'état écologique des cours d'eau. J'ai de même pu participer à une campagne sur le terrain du programme « variabilité stationnelle » qui constitue une des études en lien avec les incertitudes, et qui correspond à un protocole dérivé de celui utilisé pour le relevé IBMR.

Avec ce stage j'ai pu me rendre compte de ce en quoi consistaient l'utilisation concrète et la mise en application de la réglementation dans le domaine de l'environnement. Avec cette mission en lien avec la DCE j'ai pu me faire une idée des différentes facettes que pouvait avoir ce type de réglementation, et le travail qui en découle.

J'ai de même pu développer considérablement mes capacités d'utilisation du logiciel R en manipulant de nombreuses données, ainsi que des tableaux et fichiers Excel. J'ai au cours de ce stage pu suivre une formation MOOC sur l'utilisation de R, et obtenir la certification.

Cette mission de 6 mois au sein d'une équipe m'a permis d'apprendre à bien m'adapter dans des conditions de travail où il fallait gérer différentes étapes/missions qui se chevauchent les unes les autres, et où il est nécessaire de savoir gérer le temps de chacune pour réussir à respecter les délais impartis.

Finalement, comme précisé dans mon rapport de stage intermédiaire, je souhaitais faire mon stage de fin d'étude dans le domaine de la Recherche vu qu'il s'agit d'un milieu qui m'attire tout particulièrement.
De plus, ayant toujours été passionnée par l'Ecologie comme peut l'attester ma Licence en Biologie des Organismes et des Ecosystèmes, le sujet de stage me permettra de bien démarrer ma carrière pour une spécialisation dans le domaine de la gestion, de la protection et de la conservation de la biocénose et du biotope.

Glossaire

Aquaticité : Les coefficients d'aquaticité, caractérisent l'affinité pour l'eau des taxons pris en compte dans les approches « macrophytes ».

Biocénose : ensemble d'être vivants (espèces animales et végétales) vivant dans un lieu de vie bien défini que l'on appelle le biotope.

Bioindicateur : indicateur constitué par une espèce végétale, fongique ou animale, ou par un groupe d'espèces ou groupement végétal dont la présence ou l'état renseigne sur certaines caractéristiques écologiques (c'est-à-dire physico-chimiques, microclimatiques, biologiques et fonctionnelles) de l'environnement ou sur l'incidence de certaines pratiques.

Biotype : dans le cas de la méthode nationale basée sur l'IBMR, types de peuplements végétaux aquatiques définis par analyse statistiques sur l'ensemble des relevés sur un réseau de référence. Ils correspondent à différents types de contextes environnementaux dans les cours d'eau.

Contributivité : un taxon contributif est un taxon utilisé pour le calcul de l'IBMR, au sens de cette méthode précisée dans la norme NF T90-395. Dans le cadre de cet indice, chaque taxon contributif est affecté d'une cote spécifique CS et d'un coefficient de sténoécie E, nécessaires pour ce calcul.

DCE : Directive Cadre Européenne sur l'Eau.

EQR : Ecological Quality Ratio, rapport entre la valeur observée de l'indice d'évaluation à la valeur de référence (écart à la référence).

Eutrophe : Milieu présentant une concentration excessive en éléments nutritifs.

GIS : Groupement d'Intérêt Scientifique, le GIS « macrophytes des eaux continentales » consiste en un réseau de compétence de scientifiques créé dans le but de structurer la recherche et la réponse aux gestionnaires dans le domaine des plantes aquatiques.

IBMR : Indice Biologique Macrophytique en Rivière, indice national français se basant sur les populations macrophytique et permettant d'évaluer le statut trophique d'un cours d'eau.

Macrophyte : désigne l'ensemble des plantes aquatiques visibles à l'œil nu.

Mésotrophe : Milieu moyennement riche en éléments nutritifs

MTR : Mean Trophic Rank, indicateur biologique utilisé au Royaume-Uni se basant sur les populations macrophytiques et utilisé pour évaluer le statut trophique en rivière

Profil écologique : préferendum écologique, correspond à la niche écologique : les conditions environnementales dans lesquelles une espèce se développe.

R : logiciel libre de traitement des données et d'analyse statistiques mettant en œuvre le langage de programmation S.

Taxon : Le taxon est une unité quelconque (genre, famille, espèce, sous-espèce, etc.) des classifications hiérarchiques des êtres vivants.

Ubiquiste : Qui peut vivre partout, qui s'adapte facilement aux milieux les plus divers.

Références bibliographiques

1. irstea. *EABX - Ecosystèmes aquatiques et changements globaux*. [cited 2015; Available from: http://www.irstea.fr/epbx.
2. irstea. *Contaminants Anthropiques et Réponses des Milieux Aquatiques (Équipe CARMA)*. [cited 2015; Available from: http://www.irstea.fr/la-recherche/unites-de-recherche/eabx/contaminants-anthropiques-et-reponses-des-milieux-aquatiques.
3. eaufrance. *La directive cadre sur l'eau*. Available from: http://www.eaufrance.fr/comprendre/la-politique-publique-de-l-eau/la-directive-cadre-sur-l-eau.
4. EuropeanUnion, *Protection et gestion de l'eau (directive-cadre sur l'eau)*, 2000.
5. AQUA-REF. Available from: http://www.aquaref.fr/.
6. Dagens, N. and C. Chauvin, *L'INCERTITUDE EN HYDROBIOLOGIE*, 2014. p. 66.
7. EuropeanCommision. *The EU Water Framework Directive - integrated river basin management for Europe*. 2015; Available from: http://ec.europa.eu/environment/water/water-framework/index_en.html.
8. JournalOfficieldelaRépubliqueFrançaise, *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*, in *NOR: DEVO1001032A* 2010.
9. Legifrance, *Arrêté du 7 août 2015 modifiant l'arrêté du 25 janvier 2010 établissant le programme de surveillance de l'état des eaux en application de l'article R. 212-22 du code de l'environnement*, 2015.
10. ISO, *ISO/CEI 17025:2005(fr) Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais*, 2005.
11. Christian, C., P. Marie-Christine, and H. Jacques, *L'Indice Biologique Macrophytique en Rivière (NF T90-395). Guide méthodologique*, 2010, AFNOR. p. 122.
12. Dandelot, S., *Les Ludwigia spp. invasives du sud de la France: historique, biosystématique, biologie et écologie.*, in *Ecologie – Hydrobiologie*2004, Université Paul Cezanne Aix-Marseille III.
13. Reyjol, Y., V. Spyrato, and L. Basilic, *Bioindication : des outils pour évaluer l'état écologique des milieux aquatiques. Perspectives en vue du 2e cycle DCE – Eaux de surface continentales.*, ONEMA, Editor 2013. p. 31.
14. Christian, C., P. Marie-Christine, and H. Jacques, *La bio-indication et les indices macrophytiques, outils d'évaluation et de diagnostic de la qualité des cours d'eau*, in *Ingénieries*2008. p. 18.
15. Haury, J., et al., *A new method to assess water trophy and organic pollution – the Macrophyte Biological Index for Rivers (IBMR): its application to different types of river and pollution.* Hydrobiologia, 2006. **vol.1**: p. p. 153-158.
16. AFNOR, *Qualité de l'eau - Guide pour l'étude des macrophytes aquatiques dans les cours d'eau*, Avril 2014
17. irstea. *Macrophytes* 2015]; Available from: http://hydrobio-dce.irstea.fr/cours-deau/macrophytes/.
18. Christian, C., L. Sandrine, and F. Thibaut, *Evaluation de l'état écologique par les macrophytes – proposition d'un système de classification. Rapport Irstea Bordeaux –ONEMA, V6*, 2013. p. 24.
19. Ministèredel'écologiedudéveloppementdurableetdel'énergie. *La Directive cadre européenne sur l'eau*. 2013 [cited 2015; Available from: http://www.developpement-durable.gouv.fr/Les-objectifs.html.
20. EuropeanCommision, *Ecological Quality Ratios for Ecological Quality Assessment in Inland and Marine Waters* 2007, Institute for Environment and Sustainability. p. 24.
21. Legifrance, *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*, 2015.
22. AIDA, *Directive n° 2000/60/CE du 23/10/00 établissant un cadre pour une politique communautaire dans le domaine de l'eau*, 2000.
23. Davey, A. and D. Garrow, *Variability components for macrophyte communities in rivers: summary report FINAL REPORT*, 2009, Environment Agency. p. 42.
24. Ryszard Staniszewski, K.S., Janina Zbierska, Jacek Lesny, Szymon Jusik & Ralph T. Clarke, *Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification.* Springer, 2006: p. 12.
25. Clarke, R.T., et al., *Derivation of a biological quality index for river sites: Comparison of the observed with the expected fauna.* Journal of Applied Statistics, 1996. **23**(2-3): p. 311-332.

26. Pentecost, A., et al., *River macrophyte sampling : methodologies and variability*. 2009: Environment Agency.
27. Wiederkehr, J., et al., *Experimental study of uncertainties on the macrophyte index (IBMR) based on species identification and cover.* Ecological Indicators, 2015: p. 242-250.
28. Wiederkehr, J., *Estimation des incertitudes associées aux indices macroinvertébrés et macrophytes pour l'évaluation de l'état écologique des cours d'eau*, in *Géochimie*2015, Université de Strasbourg. p. 212.
29. UniversitéParis-Sud. *Introduction à la statistique avec R*. 2015; Available from: https://www.france-universite-numerique-mooc.fr/courses/UPSUD/42001S03/session03/courseware/f6b07f152dd542f8a74a73336e2f3485/9f0f85032c8544bea3f60aae7166f70f/.
30. Cemagref. *phpMyAdmin*. Available from: http://base-lhq.lyon.cemagref.fr/mysql/.
31. Laplace-Treyture, C., et al., *Guide pratique de détermination des algues macroscopiques d'eau douce et de quelques organismes hétérotrophes*, C. Les éditions d'Irstea Bordeaux, Editor 2014. p. 204.
32. Rodriguez, S. and J.-P. Vergon, *Guide pratique de détermination générique des algues macroscopiques d'eau douce*. 1996: Ministère de l'environnement. 110.
33. Bailly, G., J.C. Vadam, and J.P. Vergon, *Guide pratique d'identification des bryophytes aquatiques*. 2004: Ministère de l'Ecologie et du Developpement Durable. 158.
34. Coudreuse, J., et al., *Les Bryophytes aquatiques et supra-aquatiques. Clé d'identification pour la mise en œuvre de l'Indice Biologique Macrophytique en Rivière*. 2005: Les études de l'Agence de l'Eau Adour-Garonne. 132.
35. Bailly, G. and O. Schaefer, *Guide illustré des Characées du nord-est de la France*. 2010: Conservatoire Botanique National de Franche-Comté. 96.
36. Mouronval, J.-B. and S. Baudouin, *Plantes aquatiques de Camargue et de Crau.* . 2010: Office national de la chasse et de la faune sauvage. 120.
37. Spencer-Jones, D. and M. Wade, *Aquatic plants A guide to recognition*. 1986: ICI Professional Products. 169.
38. Fare, A., A. Dutartre, and J.P. Rebillard, *Les principaux végétaux aquatiques du Sud-Ouest de la France*. 2011: Agence de l'Eau Adour-Garonne. 190.
39. Prelli, R., *Les fougères et plantes alliées de France et d'Europe occidentale*. Belin ed. 2001.
40. Birk, S., et al., *Report on the Central Baltic River GIG Macrophyte Intercalibration Exercise*, 2007. p. 82.
41. Dutartre, A., H. Jacques, and P. Marie-Christine, *Plantes aquatiques d'eau douce: biologie, écologie et gestion . Le groupement d'intérêt scientifique "Macrophytes des eaux continentales".* Ingénieries, 2008: p. 2.
42. Wasson, J., et al., *Les hydro-écorégions de France métropolitaine. Approche régionale de la typologie des eaux courantes pour la définition des peuplements de référence d'invertébrés. Rapport Cemagref – Ministère de l'écologie et du développement durable*, juin 2002. p. 73.
43. Ministèredelécologiedudéveloppementdurableetdelénergie, *Mise en œuvre de la directive-cadre sur l'eau. Pour un bon état des eaux en 2015. Document technique*, 2012. p. 35.
44. Legifrance, *LOI n° 2010-788 du 12 juillet 2010 portant engagement national pour l'environnement (1)*, 2010.

Fiche compétences

Compétence	Acquisition	Justification
Réaliser une analyse de l'état de l'Environnement dans un contexte donné et évaluer la situation	Très bonne	-Etude bibliographique (étude des incertitudes hydrobiologiques sur les macrophytes et synthèse des confusions connues) -Analyse des travaux antérieurs, en particulier celui de J. Wiederkehr -Consultation des experts du GIS « Macrophytes des eaux continentales » - Utilisation de la méthodologie nationale
Réaliser une analyse prospective des impacts environnementaux potentiels d'une activité ou d'un produit	Très bonne	-Analyse multi-critère -Collection de données -Evaluation de ces données -Anticiper, prévoir, dresser des hypothèses de travail pertinentes et identifier de nouvelles problématiques -Formulation d'hypothèses d'amélioration de la démarche -Utilisation de R -Compte-rendu du travail devant l'équipe EABX -Travail selon directive européenne -Contribution à l'identification des sources d'incertitude des méthodes d'évaluation - Estimation des impacts de la confusion taxinomique sur l'évaluation de l'état écologique - Proposition de protocoles de prise en compte de ces niveaux de confiance ou des ajustements correctifs dans les méthodes.
Concevoir des outils d'action pour gérer ou prévenir des atteintes à l'Environnement	Bonne	-Réalisation d'un travail préparatoire et exploratoire faisant partie du programme d'évaluation des incertitudes -Etude technique -Réponse aux exigences d'une directive sur l'eau -Conduite d'un travail de recherche
Assurer la maîtrise d'œuvre d'une installation de prévention ou de traitement des pollutions	Moyenne	-Respect du cahier des charges du projet -Gérer les imprévus, les dysfonctionnements, les retards dans l'exécution des ouvrages et trouver des solutions

Prendre des décisions ou aider à la décision pour gérer ou prévenir les atteintes à l'Environnement	Très bonne	-Aide à la décision pour l'application d'actions sur les cours d'eau qui en ont le plus besoin -Tenir compte des contraintes de toutes natures -Construire des scénarii pertinents, analyser avantages et inconvénients des différentes hypothèses (scénarii) -Comparer les différentes solutions possibles et mener des analyses multicritères -Vision prospective du travail
Piloter des dispositifs d'action pour prévenir ou gérer une situation de dégradation de l'Environnement	Bonne	-Etablir des processus, des procédures, et adapter des techniques et méthodologies en fonction des besoins -Faire face à des situations imprévues, à des dysfonctionnements, et leur trouver des solutions -Utiliser des logiciels professionnels (+obtention du certificat MOOC R)

Annexes

Annexe 1 : biotypes identifiés et valeurs d'IBMR de référence associées [18]

Groupes	Biotypes de cours d'eau	IBMR de référence
1	cours d'eau calcaires de taille moyenne mésotrophes courants	11,17
2	cours d'eau acides de taille moyenne mésotrophes courants	13,09
3	cours d'eau cristallins très petits à petits oligotrophes très courants	14,61
4	cours d'eau très petits à petits calcaires oligotrophes, courants à très courants	12,94
5	cours d'eau de taille grande à moyenne méso-eutrophes à eutrophes faiblement courants, berges déposées	9,38
6	cours d'eau très petits à petits acides oligotrophes très courants à *Lemanea* sp. et *Oscillatoria* sp.	14,00
7	cours d'eau petits à *Rhynchostegium riparioides* et *Brachythecium rivulare*	13,17
8	cours d'eau grands à très grands eutrophes	8,5

Annexe 2 : Onglet "Biblio" du répertoire rassemblant toutes les sources bibliographiques

	Sources bibliographiques
	Algues
1	Laplace-Treyture C., Peltre M.C., Lambert E., Rodriguez S., Vergon J.P., Chauvin C. (2014). Guide pratique de détermination des algues macroscopiques d'eau douce et de quelques organismes hétérotrophes. Les éditions d'Irstea Bordeaux, Cestas. 204 p.
2	Rodriguez S., Vergon J.-P. (1996). Guide pratique de détermination générique des algues macroscopiques d'eau douce. Ministère de l'environnement. 110p.
	Bryophytes
3	Bailly G., Vadam J.C., Vergon J.P. (2004). Guide pratique d'identification des bryophytes aquatiques. Ministère de l'Ecologie et du Developpement Durable. DIREN Franche-Comté. 158 p.
4	Coudreuse J., Haury J., Bardat J., Rebillard J.P. (2005). Les Bryophytes aquatiques et supra-aquatiques. Clé d'identification pour la mise en œuvre de l'Indice Biologique Macrophytique en Rivière. Les études de l'Agence de l'Eau Adour-Garonne. 132 p.
	Characées
5	Bailly G. et Schaefer O. (2010). Guide illustré des Characées du nord-est de la France. Conservatoire Botanique National de Franche-Comté. 96p.
6	Mouronval J.-B., Baudouin S. (2010). Plantes aquatiques de Camargue et de Crau. Office national de la chasse et de la faune sauvage. 120p.
7	Spencer-Jones D., Wade M. (1986). Aquatic plants. A guide to recognition. ISBN 0 901747 03 3. ICI Professional Products
	Phanérogames
8	Fare A., Dutartre A., Rebillard J.P. (2011). Les principaux végétaux aquatiques du Sud-Ouest de la France. Agence de l'Eau Adour-Garonne. 190 p.
6	Mouronval J.-B., Baudouin S. (2010). Plantes aquatiques de Camargue et de Crau. Office national de la chasse et de la faune sauvage. 120p.
7	Spencer-Jones D., Wade M. (1986). Aquatic plants. A guide to recognition. ISBN 0 901747 03 3. ICI Professional Products
	Ptéridophytes
9	Prelli R. (2001). Les fougères et plantes alliées de France et d'Europe occidentale. Belin. ISBN 2-7011-2802-1. 431p.
7	Spencer-Jones D., Wade M. (1986). Aquatic plants. A guide to recognition. ISBN 0 901747 03 3. ICI Professional Products
8	Fare A., Dutartre A., Rebillard J.P. (2011). Les principaux végétaux aquatiques du Sud-Ouest de la France. Agence de l'Eau Adour-Garonne. 190 p.

Characées | Phanérogames | Ptéridophytes | **Biblio**

Annexe 3 : notice explicative envoyée aux experts

Incertitudes dans les mesures « macrophytes en cours d'eau »

- *Intérêt de la matrice*

L'évaluation des sources d'incertitude dans les mesures hydrobiologiques est rendue nécessaire par les prescriptions de maitrise de la confiance dans l'évaluation de l'état écologique au sens de la surveillance en réseaux DCE. Un programme de recherche méthodologique est engagé par l'équipe CARMA (Irstea Bordeaux) sur l'évaluation expérimentale de ces incertitudes, pour les protocoles Diatomées et Macrophytes en cours d'eau.

Un des aspects étudiés est l'impact des confusions taxinomiques dans la chaine d'évaluation, pour différentes métriques et indicateurs d'état écologique. Il s'agit avec cette matrice d'estimer les erreurs susceptibles d'être faites au niveau de la détermination taxonomique par les opérateurs, lors de relevés de macrophytes réalisés selon le protocole décrit dans la norme NF T90-395. La constitution de cette matrice est la base du travail d'analyse de propagation des erreurs dans la chaine d'évaluation.

L'avis d'un panel d'experts impliqués dans la détermination de macrophytes selon ce protocole et dans la validation de déterminations réalisées par les opérateurs de réseaux de mesure DCE est sollicité, afin d'établir cette matrice de la façon la plus complète et consensuelle possible. Il s'agit d'estimer d'une part les couples d'espèces pour lesquels des confusions sont avérées, d'autre part la fréquence statistique de ces confusions dans les relevés réalisés en routine.

L'objectif n'est donc pas uniquement d'établir les risques de déterminations erronées de certaines espèces en fonction de critères purement taxinomiques (clés existantes, imprécisions nomenclaturales, etc.), mais d'intégrer également l'ensemble des paramètres qui font que le résultat final d'une mesure « macrophytes » (liste floristique assortie des abondances) peut être entaché d'erreurs dans la dénomination des taxons inventoriés.

Une liste de couples de taxons pouvant prêter à confusion a été préétablie afin de vous faire gagner du temps lors du remplissage de la matrice. Toutefois si vous estimez que des couples importants ont été omis, un onglet a été expressément établi afin que vous puissiez y ajouter ces couples.

- *Principe d'utilisation de la matrice*

  - 1 onglet par groupe (Algues, Bryophytes, Characées, Hétéroptrophes, Lichens Phanérogames, Ptéridophytes)

  - 1 onglet pour les confusions intergroupes

  - 1 onglet pour l'ajout de couples à comparer qui auraient été oubliés
  ⇨ attention à ne pas faire de faute dans l'orthographe du nom du taxon ou de double espace lors de l'ajout d'un couple

  - Remplir avec le système de classe de 0 → 3 :

    - 0 - pas de confusion possible,
    - 1 - confusion peu fréquente, possible pour certaines formes ou par des opérateurs peu familiers de ces groupes,
    - 2 - confusion assez courante dans certaines conditions, pour certaines formes ou pour des opérateurs non spécialistes de ces groupes/genres,
    - 3 - confusion très fréquente dans la plupart des cas, pour toutes les formes du taxon, pour tous les opérateurs non spécialistes de ces groupes ou des types de peuplements concernés.

  ⚠ Pour un même couple, la probabilité de confusion n'est pas forcément la même dans les 2 sens (probabilité de confondre A avec B ≠ probabilité de confondre B avec A), par exemple entre une espèce A ubiquiste à très large répartition et une autre B spécifique de certaines régions ou types particuliers de cours d'eau (B sera alors souvent dénommée A, alors que l'inverse sera très rare). Cette caractéristique peut être notée en observation.

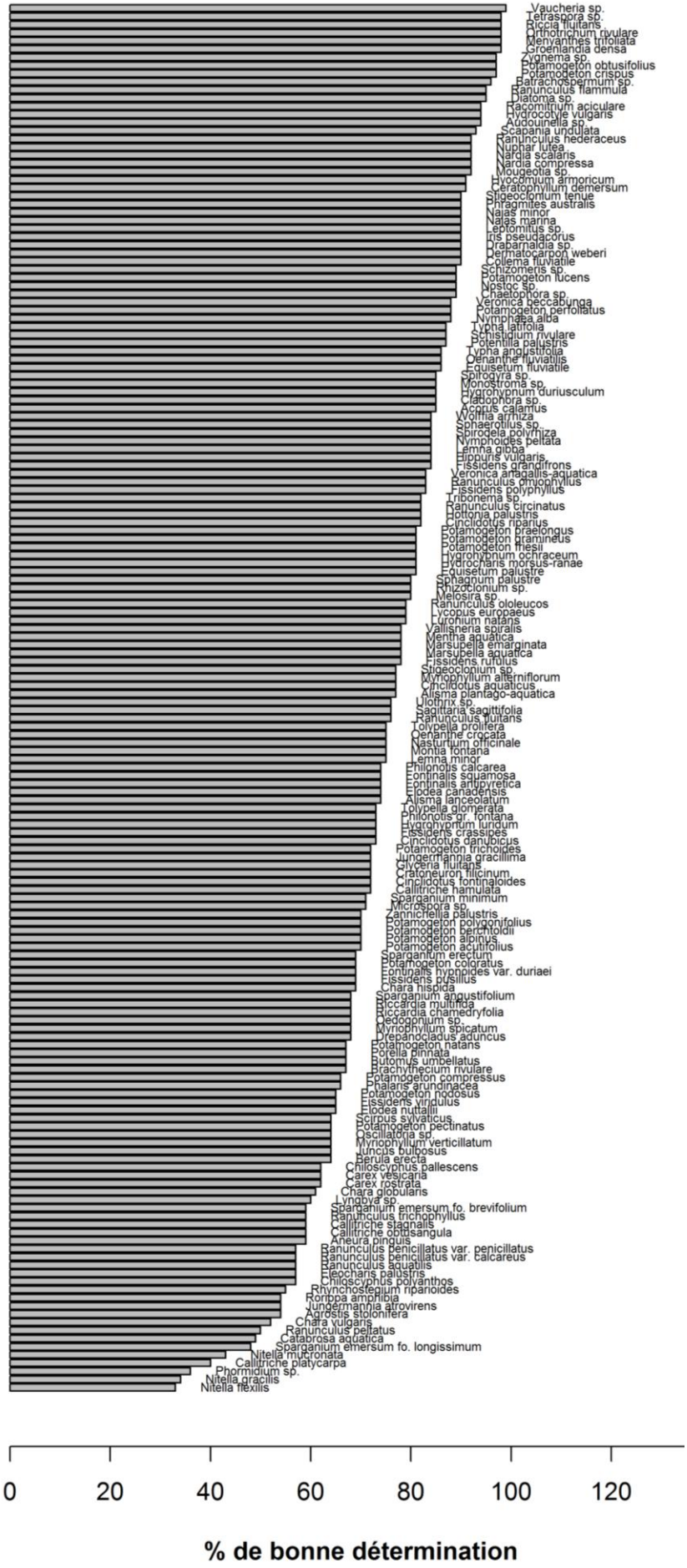
- *N.B. 1* : L'évaluation de ces confusions est considérée pour le cas d'opérateurs ayant la compétence et l'expérience dans l'identification d'espèces nécessaire et suffisante pour la mise en œuvre du protocole décrit dans la norme, mais n'étant pas spécialistes des différents groupes de macrophytes. En effet il s'agit généralement d'hydrobiologistes de bureaux d'étude. Cette identification est faite en utilisant des guides ou des ouvrages courants, et, le cas échéant, du matériel d'optique standard dans un laboratoire d'hydrobiologie.
  Il faut garder en tête qu'il s'agit d'estimer les incertitudes déjà présentes dans les relevés de mesures macrophytes réalisées dans les réseaux de mesure DCE.

- *N.B. 2* : Afin de permettre le traitement des résultats utilisant les bases de données et les référentiels méthodologiques des systèmes DCE, il convient d'utiliser le référentiel Irstea d'appellation des taxons, fourni dans le premier onglet du tableau, et de se référer à la liste des noms d'espèces retenus et de leurs synonymes.

Contact : Manon Jeanty, Irstea Bordeaux manon.jeanty@irstea.fr

Annexe 4 : Tableau de tri des taxons indicateurs par % de bonne détermination des taxons contributifs

cd_taxon	nom taxon	% de bonne détermination	cote spécifique	Type_morph
NITFLE	Nitella flexilis	33	14	Algue
NITGRA	Nitella gracilis	34	14	Algue
PHOSPX	Phormidium sp.	36	13	Algue
CALPLA	Callitriche platycarpa	40	10	Phanérogame
NITMUC	Nitella mucronata	43	14	Algue
SPAEML	Sparganium emersum fo. longissimum	48	7	Phanérogame
CATAQU	Catabrosa aquatica	49	11	Phanérogame
RANPEL	Ranunculus peltatus	50	12	Phanérogame
CHAVUL	Chara vulgaris	52	13	Algue
AGRSTO	Agrostis stolonifera	54	10	Phanérogame
JUGATR	Jungermannia atrovirens	54	19	Bryophyte
RORAMP	Rorippa amphibia	54	9	Phanérogame
RHYRIP	Rhynchostegium riparioides	55	12	Bryophyte
CHIPOL	Chiloscyphus polyanthos	57	15	Bryophyte
ELEPAL	Eleocharis palustris	57	12	Phanérogame
RANAQU	Ranunculus aquatilis	57	11	Phanérogame
RANPEC	Ranunculus penicillatus var. calcareus	57	13	Phanérogame
RANPEE	Ranunculus penicillatus var. penicillatus	57	12	Phanérogame
ANEPIN	Aneura pinguis	59	14	Bryophyte
CALOBT	Callitriche obtusangula	59	8	Phanérogame
CALSTA	Callitriche stagnalis	59	12	Phanérogame
RANTRI	Ranunculus trichophyllus	59	11	Phanérogame
SPAEMB	Sparganium emersum fo. brevifolium	59	13	Phanérogame
LYNSPX	Lyngbya sp.	60	10	Algue
CHAGLO	Chara globularis	61	13	Algue
CARROS	Carex rostrata	62	15	Phanérogame
CARVES	Carex vesicaria	62	12	Phanérogame
CHIPAL	Chiloscyphus pallescens	62	14	Bryophyte
BERERE	Berula erecta	64	14	Phanérogame
JUNBUL	Juncus bulbosus	64	16	Phanérogame
MYRVER	Myriophyllum verticillatum	64	12	Phanérogame
OSCSPX	Oscillatoria sp.	64	11	Algue
POTPEC	Potamogeton pectinatus	64	2	Phanérogame
SCISYL	Scirpus sylvaticus	64	10	Phanérogame
ELONUT	Elodea nuttallii	65	8	Phanérogame
FISVIR	Fissidens viridulus	65	11	Bryophyte
POTNOD	Potamogeton nodosus	65	4	Phanérogame
PHAARU	Phalaris arundinacea	66	10	Phanérogame
POTCOM	Potamogeton compressus	66	6	Phanérogame
BRARIV	Brachythecium rivulare	67	15	Bryophyte
BUTUMB	Butomus umbellatus	67	9	Phanérogame
PORPIN	Porella pinnata	67	12	Bryophyte
POTNAT	Potamogeton natans	67	12	Phanérogame
DREADU	Drepanocladus aduncus	68	15	Bryophyte
MYRSPI	Myriophyllum spicatum	68	8	Phanérogame
OEDSPX	Oedogonium sp.	68	6	Algue
RICCHA	Riccardia chamedryfolia	68	15	Bryophyte
RICMUL	Riccardia multifida	68	15	Bryophyte
SPAANG	Sparganium angustifolium	68	19	Phanérogame
CHAHIS	Chara hispida	69	15	Algue
FISPUS	Fissidens pusillus	69	14	Bryophyte
FONHYD	Fontinalis hypnoides var. duriaei	69	14	Bryophyte
POTCOL	Potamogeton coloratus	69	20	Phanérogame
SPAERE	Sparganium erectum	69	10	Phanérogame
POTACU	Potamogeton acutifolius	70	12	Phanérogame
POTALP	Potamogeton alpinus	70	13	Phanérogame
POTBER	Potamogeton berchtoldii	70	9	Phanérogame
POTPOL	Potamogeton polygonifolius	70	17	Phanérogame
ZANPAL	Zannichellia palustris	70	5	Phanérogame
MICSPX	Microspora sp.	71	12	Algue
SPAMIN	Sparganium minimum	71	15	Phanérogame
CALHAM	Callitriche hamulata	72	12	Phanérogame
CINFON	Cinclidotus fontinaloides	72	12	Bryophyte
CRAFIL	Cratoneuron filicinum	72	18	Bryophyte
GLYFLU	Glyceria fluitans	72	14	Phanérogame
JUGGRA	Jungermannia gracillima	72	20	Bryophyte
POTTRI	Potamogeton trichoides	72	7	Phanérogame
CINDAN	Cinclidotus danubicus	73	13	Bryophyte
FISCRA	Fissidens crassipes	73	12	Bryophyte
HYGLUR	Hygrohypnum luridum	73	19	Bryophyte
PHIFOG	Philonotis gr. fontana	73	18	Bryophyte
TOLGLO	Tolypella glomerata	73	12	Algue
ALILAN	Alisma lanceolatum	74	9	Phanérogame
ELOCAN	Elodea canadensis	74	10	Phanérogame
FONANT	Fontinalis antipyretica	74	10	Bryophyte
FONSQU	Fontinalis squamosa	74	16	Bryophyte
PHICAL	Philonotis calcarea	74	18	Bryophyte
LEMMIN	Lemna minor	75	10	Phanérogame
MONFON	Montia fontana	75	15	Phanérogame
NASOFF	Nasturtium officinale	75	11	Phanérogame
OENCRO	Oenanthe crocata	75	12	Phanérogame
TOLPRO	Tolypella prolifera	75	15	Algue
RANFLU	Ranunculus fluitans	76	10	Phanérogame
SAGSAG	Sagittaria sagittifolia	76	6	Phanérogame
ULOSPX	Ulothrix sp.	76	10	Algue

ALIPLA	Alisma plantago-aquatica	77	8	Phanérogame
CINAQU	Cinclidotus aquaticus	77	15	Bryophyte
MYRALT	Myriophyllum alterniflorum	77	13	Phanérogame
STISPX	Stigeoclonium sp.	77	13	Algue
FISRUF	Fissidens rufulus	78	14	Bryophyte
MARAQU	Marsupella aquatica	78	19	Bryophyte
MAREMA	Marsupella emarginata	78	20	Bryophyte
MENAQU	Mentha aquatica	78	12	Phanérogame
VALSPI	Vallisneria spiralis	78	8	Phanérogame
LURNAT	Luronium natans	79	14	Phanérogame
LYCEUR	Lycopus europaeus	79	11	Phanérogame
RANOLO	Ranunculus ololeucos	79	19	Phanérogame
MELSPX	Melosira sp.	80	10	Algue
RHISPX	Rhizoclonium sp.	80	4	Algue
SPHPAL	Sphagnum palustre	80	20	Bryophyte
EQUPAL	Equisetum palustre	81	10	Ptéridophyte
HYDMOR	Hydrocharis morsus-ranae	81	11	Phanérogame
HYGOCH	Hygrohypnum ochraceum	81	19	Bryophyte
POTFRI	Potamogeton friesii	81	10	Phanérogame
POTGRA	Potamogeton gramineus	81	13	Phanérogame
POTPRA	Potamogeton praelongus	81	13	Phanérogame
CINRIP	Cinclidotus riparius	82	13	Bryophyte
HOTPAL	Hottonia palustris	82	12	Phanérogame
RANCIR	Ranunculus circinatus	82	10	Phanérogame
TRISPX	Tribonema sp.	82	11	Algue
FISPOL	Fissidens polyphyllus	83	20	Bryophyte
RANOMI	Ranunculus omiophyllus	83	19	Phanérogame
VERANA	Veronica anagallis-aquatica	83	11	Phanérogame
FISGRN	Fissidens grandifrons	84	15	Bryophyte
HIPVUL	Hippuris vulgaris	84	12	Phanérogame
LEMGIB	Lemna gibba	84	5	Phanérogame
NYPPEL	Nymphoides peltata	84	10	Phanérogame
SPRPOL	Spirodela polyrhiza	84	6	Phanérogame
SPTSPX	Sphaerotilus sp.	84	0	Hétérotrophe
WOLARH	Wolffia arrhiza	84	6	Phanérogame
ACOCAL	Acorus calamus	85	7	Phanérogame
CLASPX	Cladophora sp.	85	6	Algue
HYGDUR	Hygrohypnum duriusculum	85	19	Bryophyte
MOOSPX	Monostroma sp.	85	13	Algue
SPISPX	Spirogyra sp.	85	10	Algue
EQUFLU	Equisetum fluviatile	86	12	Ptéridophyte
OENFLU	Oenanthe fluviatilis	86	10	Phanérogame
TYPANG	Typha angustifolia	86	6	Phanérogame
POEPAL	Potentilla palustris	87	16	Phanérogame
SCSRIV	Schistidium rivulare	87	15	Bryophyte
TYPLAT	Typha latifolia	87	8	Phanérogame
NYMALB	Nymphaea alba	88	12	Phanérogame
POTPER	Potamogeton perfoliatus	88	9	Phanérogame
VERBEC	Veronica beccabunga	88	10	Phanérogame
CHESPX	Chaetophora sp.	89	12	Algue
NOSSPX	Nostoc sp.	89	9	Algue
POTLUC	Potamogeton lucens	89	7	Phanérogame
SCHSPX	Schizomeris sp.	89	1	Algue
COLFLU	Collema fluviatile	90	17	Lichen
DERWEB	Dermatocarpon weberi	90	16	Lichen
DRASPX	Draparnaldia sp.	90	18	Algue
IRIPSE	Iris pseudacorus	90	10	Phanérogame
LEPSPX	Leptomitus sp.	90	0	Hétérotrophe
NAJMAR	Najas marina	90	5	Phanérogame
NAJMIN	Najas minor	90	6	Phanérogame
PHRAUS	Phragmites australis	90	9	Phanérogame
STITEN	Stigeoclonium tenue	90	1	Algue
CERDEM	Ceratophyllum demersum	91	5	Phanérogame
HYOARM	Hyocomium armoricum	91	20	Bryophyte
MOUSPX	Mougeotia sp.	92	13	Algue
NARCOM	Nardia compressa	92	20	Bryophyte
NARSCA	Nardia scalaris	92	20	Bryophyte
NUPLUT	Nuphar lutea	92	9	Phanérogame
RANHED	Ranunculus hederaceus	92	12	Phanérogame
SCAUND	Scapania undulata	93	17	Bryophyte
AUDSPX	Audouinella sp.	94	13	Algue
HYRVUL	Hydrocotyle vulgaris	94	14	Phanérogame
RACACI	Racomitrium aciculare	94	18	Bryophyte
DIASPX	Diatoma sp.	95	12	Algue
RANFLA	Ranunculus flammula	95	16	Phanérogame
BATSPX	Batrachospermum sp.	96	16	Algue
POTCRI	Potamogeton crispus	97	7	Phanérogame
POTOBT	Potamogeton obtusifolius	97	10	Phanérogame
ZYGSPX	Zygnema sp.	97	13	Algue
GRODEN	Groenlandia densa	98	11	Phanérogame
MEYTRI	Menyanthes trifoliata	98	16	Phanérogame
ORTRIV	Orthotrichum rivulare	98	15	Bryophyte
RIIFLU	Riccia fluitans	98	8	Bryophyte
TETSPX	Tetraspora sp.	98	12	Algue
VAUSPX	Vaucheria sp.	99	4	Algue

Annexe 5 : Graphique du tri des taxons indicateurs par % de bonne détermination des taxons contributifs