ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Ecole Doctorale ABIES
Agriculture Agriculture Alimentation Food Biologie Biology Environnement Environment Santé Health

Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : Ecologie

présentée et soutenue publiquement par

Anahita MARZIN

le 11 janvier 2013

Ecological assessment of running waters using bio-indicators: associated variability and uncertainty

Directeur de thèse : **Didier PONT**

Co-encadrement de la thèse : **Jérôme BELLIARD**

Jury

M. Jérôme BELLIARD, Docteur, HBAN, Irstea	Co-directeur
M. Tom BUIJSE, Doctor, Deltares	Examinateur
Mme Véronique GOURAUD, Docteur, LNHE, EDF R&D	Examinateur
M. Daniel HERING, Professor, Aquatic Ecology, University of Duisburg-Essen	Examinateur
M. Bernard HUGUENY, Directeur de recherche, UMR 7208 BOREA, MNHN	Rapporteur
M. Robert M. HUGHES, Professor, Dep. of Fisheries & Wildlife, Oregon State University	Rapporteur
M. Pierre MARMONIER, Professeur, CNRS UMR 5023 LEHNA, Université Lyon 1	Examinateur
M. Didier PONT, Directeur de recherche, HBAN, Irstea	Directeur

Irstea - Hydrosystems and Bioprocesses
1, rue Pierre-Gilles de Gennes CS 10030
92 761 Antony Cedex – France

A mon grand-père, Samad Vodjdani

بنشین بر لب جوی و گذر عمر ببین کاین اشارت ز جهان گذران ما را بس

حافظ شیرازی

« Assieds-toi au bord du ruisseau, Laisse aller ton regard au fil de l'eau,
Contemple le passage de la vie, Que ce signe d'un monde éphémère nous suffit ! »

« Beside a river seat thee on the sward, It floweth past – so flows thy life away,
So sweetly, swiftly, fleets our little day – Swift, but enough for me! »

Hâfez-e-Shirâzi

ACKNOWLEDGMENTS

Ces remerciements vont tout d'abord aux deux personnes à l'origine de ce projet : Merci donc à Didier Pont et Jérôme Belliard de m'avoir accordé leur confiance et leur soutien tout au long de cette thèse, de m'avoir laissé la liberté de chercher, de me tromper, de recommencer, de m'acharner, de changer de cap, de recommencer encore tout en me guidant de leurs précieux conseils.

J'aimerais remercier les membres de mon jury de thèse, Tom Buisje (Deltares, Utrech), Véronique Gouraud (EDF R&D, Chatou), Pierre Marmonier (CNRS, Lyon), Daniel Hering (UDE, Essen) et tout particulièrement les rapporteurs, Robert Hughes (Oregon State University, Corvallis) et Bernard Hugueny (MNHN, Paris) d'avoir accepté d'évaluer ce travail.

Merci aux membres de mon comité de thèse, Piet Verdonschot (Altera, Wageningen), Philippe Usseglio (UPMV-CNRS, Metz) et Virginie Archaimbault (Irstea, Antony), les "macroinvertebristes", Christian Chauvin (Irstea, Bordeaux), le "macrophytiste", François Delmas (Irstea, Bordeaux), le "diatomiste", Laurence Tissot, Véronique Gouraud et Cécile Delattre (EDF R&D, Châtou) ainsi qu'à Jean-Marie Mouchel (Université Paris VI, Paris), d'avoir accompagné ma réflexion au cours de ce travail de thèse, pour leurs temps, leurs conseils et remarques constructives. En particulier, merci à Christian, pour la découverte des rivières bourguignonnes et des relevés macrophytiques.

Merci à EDF R&D d'avoir contribuer au financement de cette thèse au travers de l'équipe commune EDF-Irstea « HYNES ».

Certaines personnes ont été indispensables à l'aboutissement de ce travail. Merci à Olivier pour sa patience, ses blagues, sa pédagogie et sa disponibilité pour les petits soucis de code R. Merci à Maxime pour la team 'incertitude', l'avion raté et les conseils statistiques et doctoresques. Merci à Guillaume pour les calculs SIG et l'aide Corine Land Cover.

Merci à tous mes collègues de l'équipe HEF pour l'ambiance chaleureuse et familiale, et les débats du midi à la cantine. Merci à Nicolas de m'avoir permis de sortir de mon bureau pour l'accompagner sur le terrain. Merci à Sarah et Céline, mes voisines de bureau, collègues de tricots, pour m'avoir supporté dans mes bons et mauvais jours.

Merci aux collègues d'Irstea et de l'unité HBAN qui m'ont permis de me sentir chez moi à Irstea Antony. Merci aux boxeurs de l'ASCCR, Amandine pour mes débuts, Marie-Louise, Pascal, Patrick, Elie et les autres. Aux frisbee-man and woman toujours motivés. A François, Pierre-Yves, Pierre, Florent et les autres « hydrologues sociaux ».

Merci à Nathalie Camus, Elizabeth Riant et Laurence Tanton pour leur gentillesse et leurs compétences qui nous facilitent tant les formalités administratives. Et bien sûr, merci à Roger pour son support informatique, les coups de poings du mercredi midi et la descente des boîtes de chocolat.

I had also the chance to travel and had some excellent scientific moments during the two European projects WISER and REFORM. I would like to thank especially the colleagues of the WISER WP 5.1, Christian Feld for his tenacity and rigour in leading this work package, Andreas Melcher for the "Danube Tour" and meeting with the extraordinary "Fischer Peppy", Piet for his precious advices and warm welcome in Wageningen. Thanks also to Altera students for showing me along their field experiments on Dutch river.

Merci aux amis d'ici et d'ailleurs, les Bunker, Bumper et Bacos boys'n'girls. Florette, Line, les Mours…

Merci à mes parents de m'avoir permis d'arriver jusqu'ici, ainsi qu'à Zouz et Sina pour leur soutien et leurs encouragements. A la famille Marzi-Vodj, et surtout aux sœurs Vodj et à Mamani pour la traduction et la calligraphie du poème de Hâfez.

Enfin merci à Alex pour son soutien, sa patience, les encouragements dans les moments de doute, les bons petits plats…

ABSTRACT

Sensitive biological measures of ecosystem quality are needed to assess, maintain or restore the ecological conditions of rivers. Since our understanding of these complex systems is imperfect, river management requires recognizing variability and uncertainty of bio-assessment for decision-making. Based on the analysis of national data sets (~ 1654 sites), the main goals of this work were (1) to test some of the assumptions that shape bio-indicators and (2) address the temporal variability and the uncertainty associated to prediction of reference conditions.

(1) This thesis highlights (i) the predominant role of physiographic factors in shaping biological communities in comparison to human pressures (defined at catchment, riparian corridor and reach scales), (ii) the differences in the responses of biological indicators to the different types of human pressures (water quality, hydrological, morphological degradations) and (iii) more generally, the greatest biological impacts of water quality alterations and impoundments.

(2) A Bayesian method was developed to estimate the uncertainty associated with reference condition predictions of a fish-based bio-indicator (IPR+). IPR+ predictive uncertainty was site-dependent but showed no clear trend related to the environmental gradient. By comparison, IPR+ temporal variability was lower and sensitive to an increase of human pressure intensity.

This work confirmed the advantages of multi-metric indexes based on functional metrics in comparison to compositional metrics. The different sensitivities of macrophytes, fish, diatoms and macroinvertebrates to human pressures emphasize their complementarity in assessing river ecosystems. Nevertheless, future research is needed to better understand the effects of interactions between pressures and between pressures and the environment.

Key-words: Bio-indication ▪ Rivers ▪ Fish ▪ Macroinvertebrates ▪ Benthic diatoms ▪ Macrophytes ▪ Uncertainty ▪ Bayesian modeling ▪ Inter-annual variability ▪ Environmental variability ▪ Reference condition ▪ Water Framework Directive.

RESUME COURT

Titre français : Indicateurs biologiques de la qualité écologique des cours d'eau : variabilités et incertitudes associées

Evaluer, maintenir et restaurer les conditions écologiques des rivières nécessitent des mesures du fonctionnement de leurs écosystèmes. De par leur complexité, notre compréhension de ces systèmes est imparfaite. La prise en compte des incertitudes et variabilités liées à leur évaluation est donc indispensable à la prise de décision des gestionnaires. En analysant des données nationales (~ 1654 sites), les objectifs principaux de cette thèse étaient de (1) tester certaines hypothèses intrinsèques aux bio-indicateurs et (2) d'étudier les incertitudes de l'évaluation écologique associées à la variabilité temporelle des bio-indicateurs et à la prédiction des conditions de référence.

(1) Ce travail met en évidence (i) le rôle prépondérant des facteurs environnementaux naturels dans la structuration des communautés aquatiques en comparaison des facteurs anthropiques (définis à l'échelle du bassin versant, du corridor riparien et du tronçon), (ii) les réponses contrastées des communautés aquatiques aux pressions humaines (dégradations hydro-morphologiques et de la qualité de l'eau) et (iii) plus généralement, les forts impacts des barrages et de l'altération de la qualité de l'eau sur les communautés aquatiques.

(2) Une méthode Bayésienne a été développée pour estimer les incertitudes liées à la prédiction des conditions de référence d'un indice piscicole (IPR+). Les incertitudes prédictives de l'IPR+ dépendent du site considéré mais aucune tendance claire n'a été observée. Par comparaison, la variabilité temporelle de l'IPR+ est plus faible et semble augmenter avec l'intensité des perturbations anthropiques.

Les résultats de ce travail confirment l'avantage d'indices multi-métriques basés sur des traits fonctionnels par rapport à ceux relatifs à la composition taxonomique. Les sensibilités différentes des macrophytes, poissons, diatomées et macro-invertébrés aux pressions humaines soulignent leur complémentarité pour l'évaluation des écosystèmes fluviaux. Néanmoins, de futures recherches sont nécessaires à une meilleure compréhension des effets d'interactions entre types de pressions et entre pressions humaines et environnement.

Mots clés : Bio-indication ▪ Rivières ▪ Poissons ▪ Macro-invertébrés ▪ Diatomées benthiques ▪ Macrophytes ▪ Incertitudes ▪ Modèles bayésiens ▪ Variabilité interannuelle ▪ Variabilité environnementale ▪ Condition de référence ▪ Directive Cadre sur l'Eau.

INDICATEURS BIOLOGIQUES DE LA QUALITE ECOLOGIQUE DES COURS D'EAU : VARIABILITES ET INCERTITUDES ASSOCIEES

Cette thèse s'est déroulée dans le cadre du projet européen WISER[1] et a contribué au développement et à la validation de l'indice piscicole français, l'IPR+[2] (Indice Poisson Rivière). Ces deux projets avaient tous deux pour objet l'aide à la mise en œuvre de la Directive Cadre sur l'Eau (DCE ; European Union 2000) par le développement d'outils complets d'évaluation de l'état écologique des cours d'eau européens. Cette thèse est partie prenante du programme de recherche de l'équipe commune entre Irstea et EDF R&D : « Hynes ».

Les rivières européennes et la DCE

Face à l'accroissement de la population mondiale, la ressource en eau potable, la production d'énergie renouvelable ainsi que le développement des industries et de l'agriculture sont devenus des enjeux sociétaux majeurs exacerbant la compétition entre les différents usages de l'eau (Huang & Chang 2003; Wang et al. 2003). Au cours du dernier millénaire, les rivières européennes ont été largement modifiées par l'homme (Karr et al. 1986; Petts et al. 1989) et sont aujourd'hui reconnues comme l'un des écosystèmes les plus menacés (Loh et al. 2005).

En 2000, le conseil des ministres et le parlement européen ont mis en place la DCE. Cette directive a pour but de prévenir et réduire la pollution des eaux, promouvoir son utilisation durable et améliorer l'état des écosystèmes aquatiques. Elle représente un cadre visant à harmoniser la gestion et la protection des eaux européennes (rivières, lacs, lagunes, eaux estuariennes, côtières et souterraines). Elle impose aux états membres d'évaluer la qualité écologique de leurs masses d'eau et de restaurer ou maintenir leur bon état écologique d'ici 2015 et au plus tard d'ici 2027. Outre la qualité de l'eau (ex : la concentration de l'eau en oxygène dissous, en phosphate ou en nitrate), l'évaluation de la qualité écologique des masses d'eau doit intégrer des indicateurs biologiques (ex. composition spécifique et

[1] Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery (www.wiser.eu), financé par l'Union Européenne (Programme Cadre 7, Thème 6, contrat No. 226273).

[2] Le développement de l'indice IPR+ a été financé par l'Office national de l'eau et des milieux aquatiques (ONEMA).

fonctionnelle) ainsi que des indicateurs du fonctionnement hydro-morphologique des cours d'eau (ex : la connectivité des milieux). Pour une rivière donnée, les bio-indicateurs comparent les caractéristiques actuelles de quatre communautés aquatiques (poissons, diatomées, macro-invertébrés, macrophytes) avec celles attendues en l'absence de pression humaines (l'approche des « conditions de référence » ; Bailey et al. 1998). Ils sont utilisés par la suite pour classer les masses d'eau en cinq catégories d'état écologique (très bon, bon, moyen, médiocre, mauvais).

Les enjeux actuels de la bio-indication

La bio-indication est intimement liée aux théories de l'écologie des communautés. Elle interprète la qualité d'une rivière d'après les communautés y résidant (ex. poissons, invertébrés, algues). En fonction de leurs traits d'histoire de vie et de leurs préférences écologiques, les organismes sont adaptés aux facteurs biotiques et abiotiques déterminant leur habitat dans une gamme de variation donnée (Suding et al. 2004). Une perturbation d'origine anthropique (i.e. pressions humaines) ou naturelle (ex. sécheresse, crue) faisant varier ces facteurs au-delà de cette gamme peut entrainer la modification de leur distribution et/ou de leur abondance (McCormick et al. 2001; Pont et al. 2006; Pont et al. 2007). Le bon état ou l'intégrité écologique d'un écosystème est définie par sa capacité à maintenir des communautés ayant une composition, une diversité biologique et un fonctionnement comparables aux écosystèmes « naturels » d'une même région (i.e. absence de pressions humaines) (Karr et al. 1986 ; Karr 1991; Karr & Chu 2000; Hamilton et al. 2010).

Chaque étape de la construction d'un bio-indicateur (observation des communautés, choix des métriques, définitions des conditions de référence, calcul de l'indice) joue un rôle clé dans la pertinence de l'évaluation finale des rivières et des mesures de gestion qui en découlent. Bien que le domaine de la bio-indication se soit considérablement développé avec la mise en place de DCE, de nombreux défis et polémiques restent ouverts (Hering et al. 2010).

Premièrement, de multiples approches de bio-indication ont été développées pour évaluer les cours d'eau européens et le choix des méthodes, des métriques (composition ou fonction des communautés) et des groupes biologiques sont souvent l'objet de discussions considérables (e.g. Bunn & Davies 2000).

Deuxièmement, les écosystèmes fluviaux étant très hétérogènes dans le temps et dans l'espace (Leopold et al. 1992; Gordon et al. 2004; Allan & Castillo 2007), la compréhension des facteurs naturels et anthropiques expliquant la variabilité des communautés (ex. structure fonctionnelle des communautés, composition en espèce) est essentielle à l'évaluation de l'état écologique. Les relations communautés-environnement ont été largement étudiées. Cependant, les influences respectives des facteurs naturels, anthropiques et de leurs interactions sont rarement différenciées.

Afin de prendre en compte la forte variabilité naturelle des écosystèmes fluviaux, deux méthodes sont principalement utilisées : l'approche « typologique » et l'approche « prédictive » (i.e. modélisation statistique des caractéristiques des communautés en fonction de l'environnement et en absence de pressions humaines). Les conditions de référence sont la base de l'évaluation. Dans les deux cas, leur définition simplifie une réalité complexe et repose sur des observations limitées. Il en résulte une incertitude autour de la condition de référence et par conséquent de l'évaluation finale. Plus généralement, de multiples sources d'incertitude peuvent jouer sur la précision de l'évaluation finale (Clarke & Hering 2006). Leur estimation et leur prise en compte lors de l'évaluation et de la prise de décision liées à la gestion des rivières européennes est un enjeu majeur (Hering et al. 2010).

Enfin, la réponse des communautés aux mesures d'amélioration des rivières est souvent imprédictible (Suding et al. 2004). L'intégration des résultats des mesures de restauration est essentielle à l'amélioration de la gestion des rivières (Palmer et al. 2007).

Objectifs de la thèse

Cette thèse aborde différents aspects et limites de l'utilisation de bio-indicateurs pour l'évaluation écologique des rivières. Des solutions et des recommandations sont proposées pour les futurs développements et utilisations des indicateurs biologiques.

Trois objectifs ont structuré ce travail :

(1) Fournir des éléments pour la sélection des groupes biologiques et des descripteurs des caractéristiques des communautés pour évaluer les cours d'eau (***P1***[3] et ***P2***).

(2) Dans le cadre du développement de l'IPR+, proposer une méthode permettant la prise en compte des incertitudes liées à la prédiction des conditions de référence (***P3***) et aborder deux sources possibles d'incertitude de l'évaluation : l'incertitude prédictive et la variabilité temporelle des bio-indicateurs (***P3*** et ***P4***).

(3) Au travers de l'exemple de l'effacement de seuil, faire le point sur les effets connus des mesures de restauration et les futurs besoins des gestionnaires (***P5***).

[3] P1 à P5 font référence aux articles présentés en fin de manuscrit

Principaux résultats et discussion générale

Quels indicateurs des impacts des pressions humaines sur les écosystèmes fluviaux ?

Les analyses menées dans ***P1***, ***P3*** et ***P4*** ont montré que les réponses des indicateurs aux pressions humaines varient fortement d'un type d'indicateur à un autre. Ces résultats confirment l'importance d'une sélection des métriques entrant dans le calcul des indicateurs basée non seulement sur le groupe biologique mais également sur la nature des métriques. Trois types d'indicateurs ont été comparés : les métriques fonctionnelles (ex. le nombre d'espèces rhéophiles), les métriques compositionnelles (ex. le nombre de juvéniles de truite) et les indicateurs multi-métriques (ex. IPR+). Les métriques fonctionnelles sont généralement celles permettant de détecter les niveaux de pressions les plus faibles alors que les indices multi-métriques répondent le plus fortement à l'ensemble des pressions humaines considérées. De plus, la seule métrique compositionnelle entrant dans le calcul de l'IPR+ (l'abondance relative des juvéniles de truite) est en moyenne moins stable dans le temps que les métriques fonctionnelles (***P4***). Ces résultats sont en adéquation avec des études antérieures suggérant que la structure fonctionnelle des communautés est moins variable dans le temps et dans l'espace que la composition taxonomique de celles-ci (Bêche et al. 2006; Fransen et al. 2011).

D'autre auteurs supposent que les traits fonctionnels sont mieux adaptés à des approches larges échelles (Statzner et al. 2001; Lamouroux et al. 2002 ; Dolédec et al. 2006). En outre, les incertitudes de prédiction de chacune des métriques étaient globalement plus fortes que celles de l'indice multi-métriques IPR+.

Ces résultats confirment, pour l'évaluation des rivières, l'avantage des indices multi-métriques combinant des métriques basées sur des traits fonctionnels écologiques et biologiques des espèces des communautés. De plus, ***P3*** illustre l'intérêt de la sélection, pour chaque site, des métriques les plus dégradées. Ainsi, les métriques les plus sensibles aux conditions de pressions sont utilisées pour refléter l'état écologique de l'écosystème.

Comment prendre en compte les imperfections de nos données, de nos connaissances et de nos méthodes ?

Les analyses des articles ***P1*** et ***P2*** ont confirmé le rôle prépondérant de la **variabilité naturelle** des facteurs abiotiques (ex. température, précipitation, géologie, taille du cours d'eau) dans la structuration des communautés aquatiques.

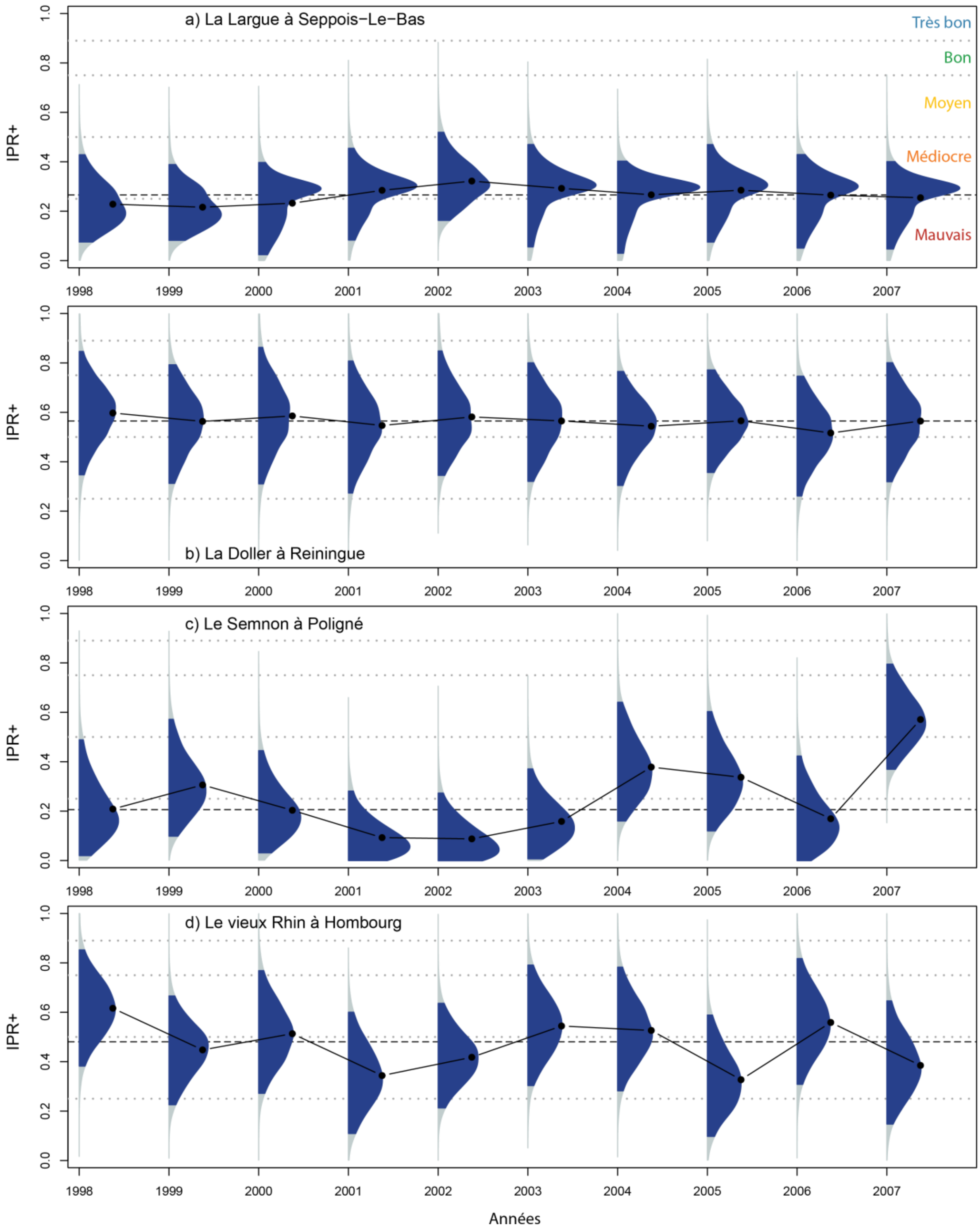


Figure 1. Variabilité interannuelle et incertitude prédictive des notes IPR+ pour quatre sites. Les fonctions de densité de probabilité de l'IPR+ sont représentées de 1998 à 2007. Zone bleue, 95% de la distribution. Zone grise, les 5% restant. Ligne noire pointillée, médiane des notes IPR+ pour les 10 ans. Point noir, note IPR+ (i.e. la moyenne de la distribution). Les lignes grises séparent les cinq classes d'état écologique.

De plus, l'analyse de la **variabilité interannuelle** des communautés piscicoles en l'absence de changement des pressions humaines illustre l'importance des processus de dynamique des populations ainsi que les effets des perturbations d'origine naturelle (ex. sécheresse ou crue) sur la composition des communautés.

Cette variabilité (bruit) doit être différenciée de l'effet des changements induits par l'homme (signal) afin de ne pas biaiser l'évaluation. Comme évoqué précédemment, l'utilisation de modèles statistiques pour prédire les caractéristiques des communautés en l'absence des pressions humaines (Oberdorff et al. 2001; Pont et al. 2006; 2007) semble appropriée à la prise en compte de cette part de la variabilité et à son élimination du signal (***P2***, ***P3***). L'effet des variations climatiques peut être pris en compte par l'intégration aux modèles de variables telles que les températures et précipitations annuelles (Linke et al. 1999; Mazor et al. 2009).

Cependant, les modèles statistiques sont par définition une simplification de la réalité. La représentativité des échantillons dépend fortement de l'effort d'échantillonnage qui peut varier d'un site à l'autre et d'une année sur l'autre. Pour ces deux raisons, la prise en compte de la variabilité naturelle ne peut être que partielle et peut biaiser la prédiction des métriques en l'absence de pressions humaines et par conséquent l'évaluation donnée par l'indice multi-métrique (i.e. incertitude de prédiction). **Ces deux sources d'incertitudes représentent un risque potentiel pour les décideurs et nécessitent d'être estimées.**

Incertitude de prédiction des conditions de référence - La méthode Bayésienne développée semble appropriée à la considération de l'incertitude prédictive d'un indice multi-métrique (Fig. 1 ; ***P3*** et ***P4***). Les notes IPR+ sont indépendantes de la variabilité naturelle et sont sensibles aux perturbations anthropiques. En particulier, l'indice IPR+ est sensible à l'augmentation du nombre de types de pressions incluant, par exemple, une modification du fonctionnement hydrologique du cours d'eau.

Les tronçons de rivière non perturbés par les activités humaines sont très rares à l'aval des cours d'eau français. Afin de pouvoir modéliser les caractéristiques des communautés de référence, les sites faiblement à modérément dégradés par les activités humaines ont donc été intégrés pour construire les modèles (***P3***). En effet, différentes définitions des conditions de référence sont utilisées en bio-indication (Stoddard et al. 2006). Alors qu'elles sont définies *stricto sensu* par l'absence de pressions humaines, certains auteurs (Pardo et al. 2012) recommandent une définition plus large permettant d'étendre l'utilisation des bio-indicateurs en zone aval des cours d'eau. **Cette réflexion met en avant la subjectivité des conditions de référence et le besoin d'établir des critères de définition communs.**

La variabilité temporelle - Pour 183 sites français présentant des pressions humaines constantes de 1998 à 2007, les évaluations données par la note IPR+ sont pour la plupart cohérentes (Fig. 1 ; ***P4***). Toutefois, la variabilité interannuelle de la note IPR+ est plus forte pour les sites très perturbés que pour des sites moins exposés à des pressions humaines. Cet effet est certainement à relier à la variabilité interannuelle des communautés piscicoles augmentant avec le niveau de pressions humaines (section 2.2 ; Fore et al. 1994; Ross et al. 1985; Collier 2008). Schaeffer et al. (2012) ont suggéré que la redondance des traits fonctionnels d'une communauté pourrait atténuer l'effet des pressions humaines sur le fonctionnement de l'écosystème. Suite à des perturbations anthropiques, la perte de certaines espèces peut entrainer une réduction de ces forces compensatrices et par conséquent déstabiliser la structure de la communauté (Franssen et al. 2011). Si comme suggéré par ces résultats, on admet que les **pressions humaines déstabilisent la structure des communautés fluviales**, la variabilité interannuelle de la note de l'indice pourrait être utilisée comme une indication supplémentaire de la dégradation de l'intégrité du système. Dans le cas de l'IPR+, l'analyse comparative de l'incertitude prédictive et de la variabilité interannuelle ont montré que la première est généralement plus forte que la deuxième (***P4***).

Quels groupes biologiques pour l'évaluation de l'intégrité des rivières ?

L'étude comparative des impacts de pressions humaines sur quatre groupes biologiques (poissons, diatomées benthiques, macrophytes, macro-invertébrés ; ***P1*** et ***P2***) a montré que selon leurs traits d'histoire de vie, les organismes et les bio-indicateurs associés répondent différemment en terme d'intensité (i.e. quelle est l'ampleur de l'impact ?) et de sensibilité (i.e. à partir de quel niveau de pression un impact est-il observé ?). Les indicateurs basés sur les diatomées et les macro-invertébrés semblent les plus sensibles à la dégradation de l'ensemble des conditions du cours d'eau (***P1***, Fig. 2). Les indicateurs diatomiques et macrophytiques sont les plus sensibles à une dégradation de la qualité de l'eau. Les indicateurs piscicoles montrent les intensités de réponse les plus fortes mais généralement une sensibilité faible. Ces différences de sensibilité peuvent être en partie expliquées par la plus grande longévité et mobilité des poissons leur permettant d'éviter et de survivre aux perturbations contrairement à des organismes sédentaires ayant des cycles de vie plus courts. De plus, les indicateurs piscicoles sont les plus impactés par les modifications hydro-morphologiques.

Ces résultats invalident en partie la règle du « one-out-all-out » (i.e. principe de l'élément le plus déclassant) supposant que chaque groupe biologique répond à un type de pressions en particulier. En effet, l'ensemble des groupes biologiques semble répondre aux différents type de pressions mais avec des intensités et des sensibilités différentes.

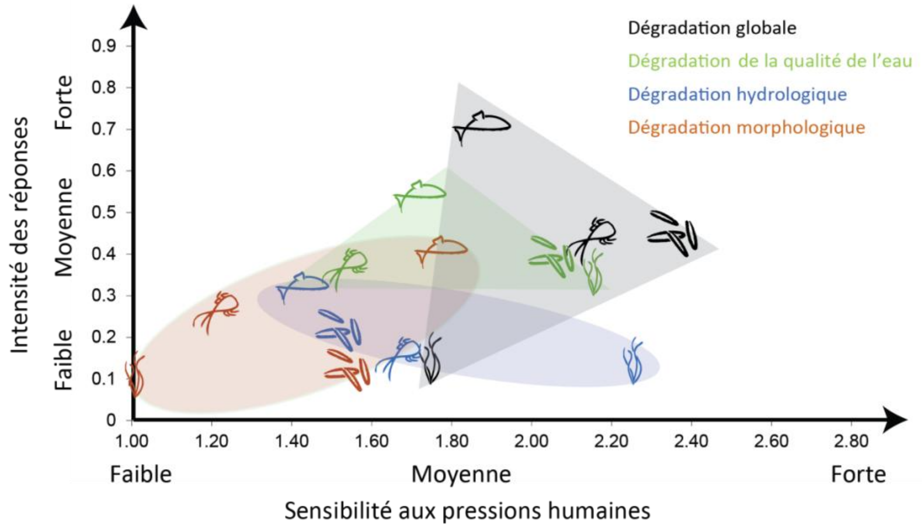


Figure 2. Les réponses médianes des métriques testées par groups biologiques (macrophytes, diatomées benthiques, poisons et macro-invertébrés). Les réponses à une perturbation globale du cours d'eau ainsi qu'à trois types de perturbation (dégradation de la qualité de l'eau, hydrologique et morphologique) sont décrites en termes d'intensité et de sensibilité.

Il est donc plus probable que certains groupes détectent les impacts des dégradations les plus faibles alors que d'autres détectent les impacts les plus sévères. **D'après ces résultats, les groupes biologiques apportent des informations complémentaires sur le niveau de dégradation des écosystèmes. L'association d'indicateurs basés sur ces différents groupes renforcerait donc la qualité de l'évaluation de l'état écologique.**

Quels sont les impacts des pressions humaines détectés par l'observation des communautés aquatiques ?

Le pourcentage du bassin versant et du corridor riparien occupés par la forêt semble être une bonne indication de la qualité globale de l'eau en amont du site. Néanmoins, des mesures à l'échelle d'un tronçon de cours d'eau sont indispensables à la description des dégradations locales de la qualité de l'eau et des processus hydro-morphologiques (***P2***).

Les analyses partielles de redondance de la variabilité de la composition des communautés piscicoles et macro-invertébrées (***P2*** ; Fig. 3) ainsi que la comparaison des impacts des pressions humaines sur les quatre groupes biologiques (poissons, macrophytes, diatomées benthiques, macro-invertébrés ; ***P1***) ont montré que **les pressions humaines agissent de manière hiérarchisée sur les communautés aquatiques**.

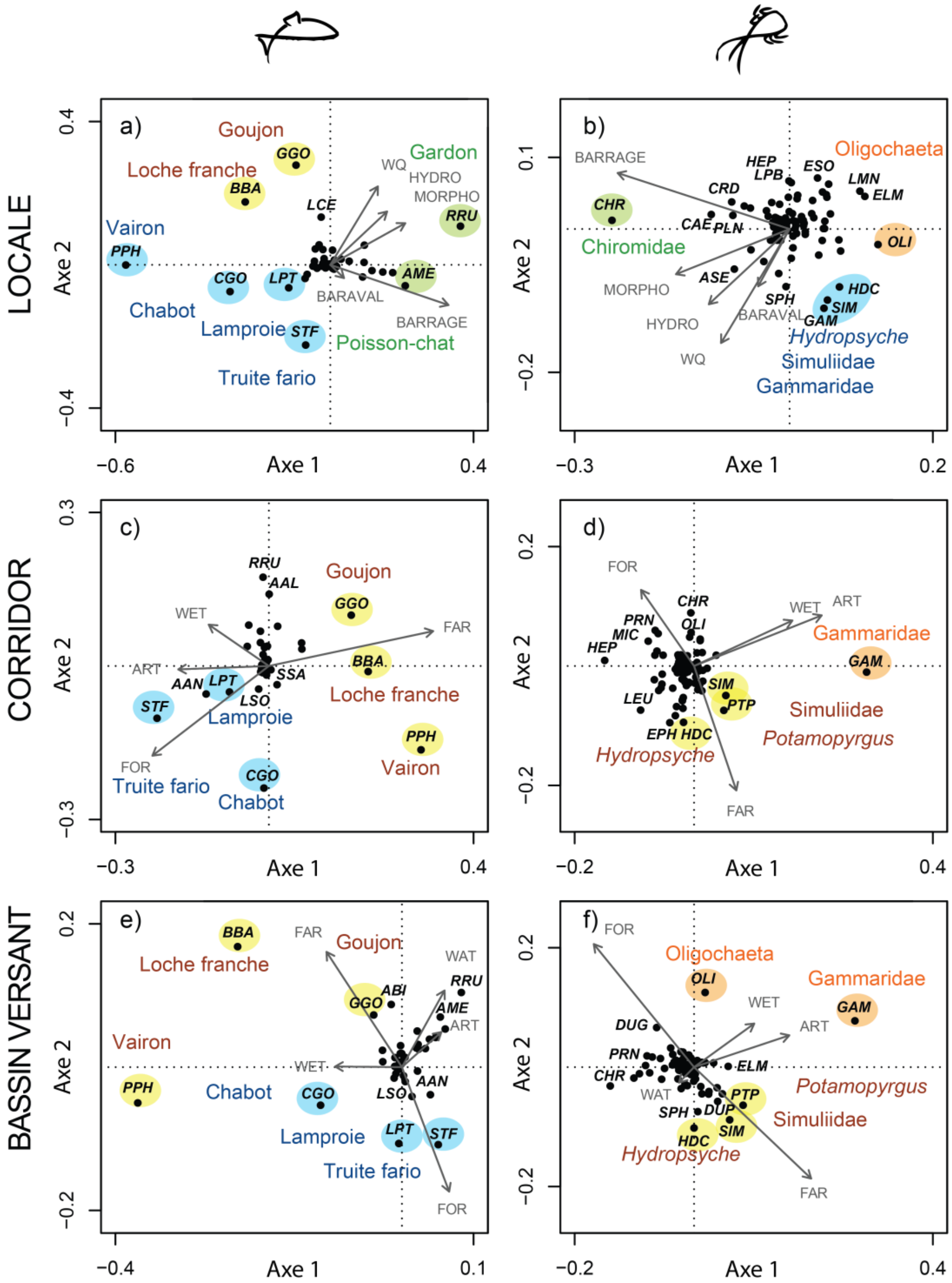


Figure 3. Biplots des six analyses partielles de redondance (pRDA). Après avoir pris en compte l'effet de l'environnement, les relations entre la composition des communautés piscicoles et macro-invertébrés et les pressions humaines sont étudiées à trois échelles :

- locale, i.e. tronçon de cours d'eau (a,b), WQ = dégradation de la qualité de l'eau, Bardown = présence d'une barrière à l'aval.
- du corridor ripарien (c, d), % d'occupation du sol du type ART = artificielle, FAR = agricole, FOR = forêt, WAT = surface en eau, WET = zones humides.
- et du bassin versant (e, f).

La longueur des flèches correspond à l'importance des relations entre les variables et les axes. Les codes de trois lettres correspondent aux taxons de poissons (ABI, *Alburnoides bipunctatus*; AAL, *Alburnus alburnus*; AME, *Ameiurus melas*; AAN, *Anguilla anguilla*; BBA, *Barbatula barbatula*; CGO, *Cottus gobio*; GGO, *Gobio gobio*; LPT, *Lampetra* sp.; PPH, *Phoxinus phoxinus*; RRU, *Rutilus rutilus*; SSA, *Salmo salar*; STF, *Salmo trutta fario*; LCE, *Squalius cephalus*; LSO, *Telestes souffia*) et de macro-invertébrés (ASE, *Asellidae* Gen. sp.; CAE, *Caenis* sp.; CHR, Chironomidae Gen. sp.; CRD, *Corixidae* Gen. sp.; DUG, *Dugesiidae* Gen. sp.; DUP, *Dupophilus* sp. Ad.; ELM, *Elmis* sp. Ad.; EPH, *Ephemera* sp.; ESO, *Esolus* sp. Ad.; GAM, *Gammaridae* Gen. sp.; HEP, *Heptageniidae* Gen. sp.; HDC, *Hydropsyche* sp.; LPB, *Leptophlebiidae* Gen. sp.; LEU, *Leuctridae* Gen. sp.; LMN, *Limnius* sp. Ad.; MIC, *Micrasema* sp.; OLI, *Oligochaeta* Gen. sp.; PLN, *Planorbidae* Gen. sp.; PTP, *Potamopyrgus* sp.; PRN, *Protonemura* sp.; SIM, *Simuliidae* Gen. sp.; SPH, *Sphaerium* sp.).

En accord avec des études antérieures (ex. Hering et al. 2006), l'accumulation de différents types de dégradations semble entrainer les plus forts impacts sur la biocénose aquatique (***P1***, ***P2***, ***P3***).

La dégradation de la qualité de l'eau et la présence d'un barrage à l'aval du tronçon sont les deux types de pressions entraînant les modifications des communautés les plus importantes (***P1***, ***P2*** ; Fig. 3). Cette hiérarchie suppose que la faune et la flore aquatiques sont très sensibles à la qualité de l'eau et qu'une augmentation de celle-ci améliorerait également l'état écologique du cours d'eau. Inversement, des mesures visant à améliorer les conditions hydro-morphologiques, sans résoudre les problèmes de qualité d'eau, risquent de ne pas avoir les effets positifs souhaités sur l'état écologique.

Les effets d'interaction entre facteurs naturels et pressions humaines et entre les différents types de pressions humaines

Les résultats des analyses menées dans les articles ***P1*** et ***P2*** suggèrent des interactions non négligeables entre différents types de pressions ainsi qu'entre variabilité naturelle et pressions anthropique. Le premier type d'interaction pose la question suivante : les effets de deux pressions différentes sur les communautés sont-ils additifs, multiplicatifs ou antagonistes ? Dans le cas très commun d'une rivière « multi-impactée », l'existence d'effets d'interaction entre pressions rend difficile le diagnostic de l'état écologique et la prédiction des impacts des pressions. Malheureusement, ces effets restent jusqu'à présent peu étudiés et très mal connus (Pont et al. 2007).

De plus, l'existence d'effets d'interaction entre l'environnement « naturel » et les pressions humaines signifierait que l'impact d'une même pression serait différent selon le type de cours d'eau concerné. Dans ce cas, l'approche de modélisation « prédictive » ne permet d'éliminer qu'une partie de l'effet de l'environnement sur la variabilité des communautés. S'il est important, cet effet complique considérablement l'interprétation de l'analyse des impacts anthropiques.

Quelle est la réponse des communautés aquatiques aux opérations de restauration ?

L'analyse de la littérature concernant les effets des effacements de seuil sur la faune et la flore aquatiques (***P5***) a souligné l'importance d'un suivi à long terme afin d'améliorer la compréhension des résultats des mesures de restauration. Alors que des effets positifs immédiats sont généralement attendus, le rétablissement du fonctionnement de l'écosystème peut prendre des années. De plus, il semble que l'efficacité des actions de restauration est rarement quantifiée et que les critères de leurs succès restent à définir. Enfin, les modèles conceptuels facilitent le regroupement et la structuration des connaissances acquises. Ils représentent un cadre de travail intéressant permettant l'identification des lacunes actuelles et des pistes à considérer pour de futures recherches.

Perspectives

Une voie d'amélioration possible des bio-indicateurs, et en particulier ceux basés sur les macrophytes, réside dans la recherche de traits fonctionnels biologiques et écologiques pertinents pour l'évaluation de l'intégrité des communautés. Par exemple, définir des **métriques et indicateurs basés sur l'étude des réseaux trophiques** pourrait constituer une approche quantitative des impacts anthropiques sur le fonctionnement des écosystèmes en considérant le rôle complexe des interactions trophiques (Friberg et al. 2011; Thompson et al. 2012).

Dans un contexte de changement global, il est indispensable de prendre en compte des données climatiques (ex. précipitations, températures) dans les modèles de prédiction des conditions de référence afin de considérer leurs effets sur les communautés (Nichols et al. 2010). De cette manière, l'évolution des conditions de référence pourra être prédite et prise en compte pour l'évaluation des rivières dans un futur proche. Cependant, d'ici 30 à 50 ans, les métriques sélectionnées pour des conditions actuelles ne seront probablement plus représentatives des écosystèmes des rivières françaises (ex. changement du pool régional d'espèces ; Logez & Pont 2012). Par conséquent, **dans le futur, les indicateurs et les conditions de référence devront être redéfinis afin de pouvoir évaluer l'état écologique des écosystèmes aquatiques.**

Une étude approfondie des effets d'interactions entre pressions humaines et environnement et entre types de pressions apparait indispensable pour améliorer la compréhension des écosystèmes fluviaux et leur gestion. Tester de tels effets suppose de très grands jeux de données permettant de recréer des plans d'expérience équilibrés. Cet aspect pourrait être abordé en premier lieu au travers de cas simples pour lesquels les effets des pressions considérées sont bien connus et au moins en partie interprétables. Par exemple, dans le cas de la présence d'un barrage et de la dégradation de la qualité de l'eau, la question de l'existence d'effets d'interaction pourrait se résumer par : les effets d'une diminution de la qualité de l'eau sur les communautés sont-ils les mêmes dans un tronçon avec ou sans barrage à l'aval ?

La méthode Bayésienne développée dans cette thèse, pourrait être utilisée pour estimer l'incertitude associée à la prédiction des conditions de référence pour d'autre bio-indicateurs. De mon point de vue, la prise en compte de l'incertitude liée à l'évaluation de la qualité écologique est indispensable à la prise de décision mais c'est aussi **l'étape préliminaire à la réduction de l'incertitude et à l'amélioration de la fiabilité des indices**.

Dans le cas de l'IPR+, l'incertitude prédictive est plus forte que la variabilité interannuelle. Sa réduction permettrait d'améliorer la fiabilité de l'indice et son pouvoir de détection des impacts des pressions. La qualité des données utilisées pour l'ajustement des modèles (effort et méthode d'échantillonnage, précision des variables environnementales, évaluation des pressions humaines) jouent un rôle important dans la maîtrise des incertitudes. Plusieurs études sont actuellement en cours pour l'évaluation des incertitudes liées aux méthodes d'échantillonnage françaises notamment pour les macro-invertébrés (Virginie Archaimbault, communication personnelle) et les poissons (Tomanova et al., en révision).

Des connaissances sur les relations entre métriques et variables environnementales, acquises par des jeux de données indépendants mais comparables, permettraient de maximiser l'utilisation des méthodes bayésiennes et de réduire l'incertitude des prédictions. De plus, des méthodes Bayésiennes plus sophistiquées permettant la réduction des incertitudes pourraient être explorées (modèles hiérarchiques ; Gelman et al. 2004). Enfin, la variabilité temporelle et l'incertitude prédictive de l'IPR+ pourraient être réduites en considérant comme critères de sélection des métriques, l'incertitude prédictive et la variabilité interannuelle.

P1, ***P2*** et ***P3*** ont montré que les bio-indicateurs développés répondent aux pressions humaines. Logiquement, l'étape suivante est l'analyse des réponses des communautés à la réduction des pressions (restauration). En effet, il n'est pas certain que l'effet des mesures de restauration soit simplement l'inverse des effets des dégradations (Moerke et al. 2004; Feld et al. 2011). Comme illustré par le cas de l'effacement de seuil, alors que la qualité de l'eau et la connectivité hydraulique sont immédiatement rétablies, les cours d'eau peuvent mettre des années à recouvrer leurs caractéristiques morphologiques. Par conséquent, les caractéristiques des systèmes après suppression des pressions ne sont pas toujours prévisibles. Suding et al. (2004) suggèrent que les échecs des mesures de restauration sont généralement liés à un manque de considération de certains paramètres ayant été modifiés : occupation du sol, perte de la connectivité, introduction d'espèces exotiques invasives...

Cette vision dynamique des écosystèmes remet en question la définition d'une situation de référence unique au cours du temps pour un type de cours d'eau donné (Hobbs & Harris 2001; Jähnig et al. 2011).

TABLE OF CONTENTS

PART I: SYNTHESIS

ABBREVIATIONS

ANOVA: Analysis of variance

BQE: Biological Quality Elements

DPSIR: Driver, Pressure, State, Impact, Response

DREAL: Direction régional de l'Environnement, de l'Aménagement et du Logement

EEA: European Environmental Agency

EQR: Ecological Quality Ratio

EU: European Union

IBI: Index of Biotic Integrity

IPR+: Indice Poisson Rivière +

OECD: Organisation for Economic Co-operation and Development

ONEMA: Office National de l'Eau et des Milieux Aquatiques ;

PCA: Principal Component Analysis

PDF: Probability Density Function

RCA: Reference Condition Approach

RDA: Redundancy Analysis

SD_U: Predictive uncertainty

SD_T: Inter-annual variability

WFD: Water Framework Directive

WISER: Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery

1. INTRODUCTION

This thesis was part of the European WISER project[1] and participated in the development and validation of the French fish-based index IPR+[2] (Indice Poisson Rivière). These two projects aimed to support the implementation of the Water Framework Directive (WFD; European Union 2000) by developing tools for the integrated assessment of the ecological status of rivers.

In this context, the main goals were to explore and test various aspects and assumptions of river biological assessment, address some limits of biological indicators and suggest improvements. To achieve these objectives, a large-scale database was compiled from the French national monitoring programs (French National Agency for Water and Aquatic Environments (ONEMA), the French water agencies (Agences de l'eau) and the French Regional Direction of Environmental Services (DREAL)), including information on river biological quality elements (BQEs, i.e. fish, macrophyte, macroinvertebrate and diatom), hydro-morphological human-induced modifications, land use and water quality (Fig. 1).

This thesis was part of the collaborative project between IRSTEA and the French Electric Company (EDF R&D): HYNES.

1.1 FROM SOCIETAL NEEDS AND ECOLOGICAL CONCEPTS TO RIVER MANAGEMENT

1.1.1 European rivers and the Water Framework Directive

"For most of the twentieth century, society remained largely unaware of the collapse of aquatic ecosystems because we saw water narrowly, as a fluid to be consumed or used as a raw material in agriculture or industry."
Karr 2006

Since humans have populated the Earth, they have relied on rivers for food, drinking water, waste removal, commerce, transportation and recreation. As the world population continues to grow, agriculture, industrialization and urbanization have increased threatening water supplies, and competition between river users exacerbated (Huang & Chang 2003; Wang et al. 2003).

[1] Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery (www.wiser.eu), funded by the European Union under the 7th Framework Programme, Theme 6, contract No. 226273.
[2] The development of the IPR+ index was supported by the French National Agency for Water and Aquatic Environments (ONEMA).

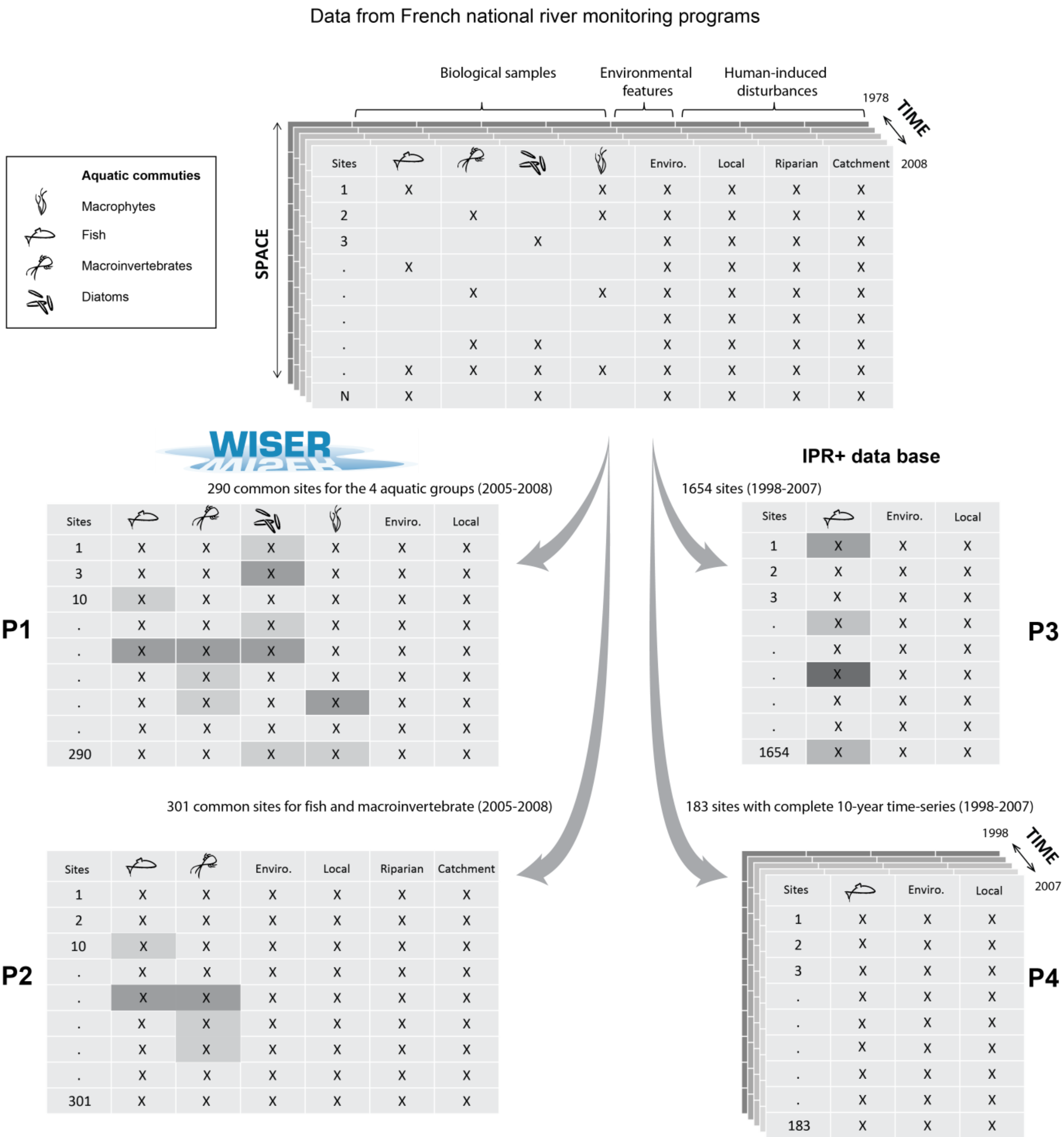


Figure 1. Description of the data sets compiled and used for the PhD thesis. Environmental features correspond to ten environmental variables assumed to be quasi-independent of human activity: catchment area (km²), catchment geological type, altitude (m), distance to the source (km), mean slope (‰), stream power (kg.m.s^{-3}), hydrological regime, mean annual air temperature (°C), mean annual air temperature amplitude (°C) and ecoregions.

Human-induced disturbances were described at three spatial scales:

- catchment and riparian corridor land use/covers (relative covers %: farming (FAR), urban and artificial (ART), forest and semi-natural (FOR), wetland (WET), water (WAT); Corine Land Cover 2000)
- reach-scale human-induced modifications of the river hydrology (hydrological regime modified, hydropeaking, water abstraction), morphology (riparian vegetation modified, artificial embankment, instream habitat modified, channel form modified, cross-section modified, diked, sedimentation) and water quality (oxygen saturation (%), BOD5 (mg O2/l), nitrite (mg NO_2/l), nitrate (mg NO_3/l), ammonia (mg NH4/l), orthophosphate (µg PO_4/l), total phosphate (µg P/l)).

P1 to ***P4***, the four scientific articles produced from this data.

Yet, rivers have experienced a long history of anthropogenic modifications of water quality, habitat structure, flow regime, energy source and biotic interaction (Karr et al. 1986; Petts et al. 1989) and have become one of the most threatened ecosystems (Loh et al. 2005).

To respond to these critical societal concerns, the Council of Ministers and the European Parliament with the contribution of non-governmental organizations finalized the WFD (European Union 2000). The WFD aims to harmonize existing European water policy and to improve the quality of European aquatic environments (rivers, lakes, transitional, coastal and ground waters). It stipulates that the EU Member States must assess the quality of their water bodies and maintain or restore their "good status" by 2015, or at the latest, 2027. The novelty of the WFD is that, in addition to water quality (e.g. water concentrations in dissolved oxygen, phosphate or nitrate), water body quality status assessment has to integrate information on the integrity of river hydro-morphological processes (e.g. river connectivity) and communities' functioning (e.g. species and functional composition) (Fig. 2). This directive is consistent with the "New Public Management" (Boston 1996) and its need for indicators to evaluate the efficiency of public policies (objectives, costs and performance) (Bouleau & Pont, submitted). The characteristics of four aquatic assemblages (diatom, fish, macroinvertebrate, macrophyte) should be compared to these communities' characteristics in absence of human-induced disturbances (the Reference Condition Approach; RCA; Bailey et al. 1998; Fig. 2) through bio-indicators. Using these indicators, a river site should be classified into one of five levels (high, good, moderate, poor, and bad status) using the "one-out-all-out rule" (Fig. 2). This rule corresponds to the precautionary principle, i.e. the BQE with the worst ecological quality is used to determine the ecological class. However, there is still a lively debate about whether it really gives the best indication of overall water body quality and the consequences on assessment results (e.g. Cartensen 2007; Hering et al. 2010). In addition, a conceptual framework (Driving force, Pressure, State, Impact, Response; DPSIR, Fig. 2) was developed to analyse cost–benefit relationships and compare efficiency of measures aiming at reducing human impacts on aquatic ecosystems (OECD 1993; 1994; EEA 1995). As a consequence, water managers needed efficient and robust tools to assess the impacts of human activities on aquatic ecosystems and meet the ambitious objectives of the WFD. European Member States had to develop their own national assessment system for each of the four BQEs. As performance of ecological assessments varies between European States, the WFD implemented an inter-calibration exercise to harmonize the understanding of "good ecological status" and the results of the different national assessment methods.

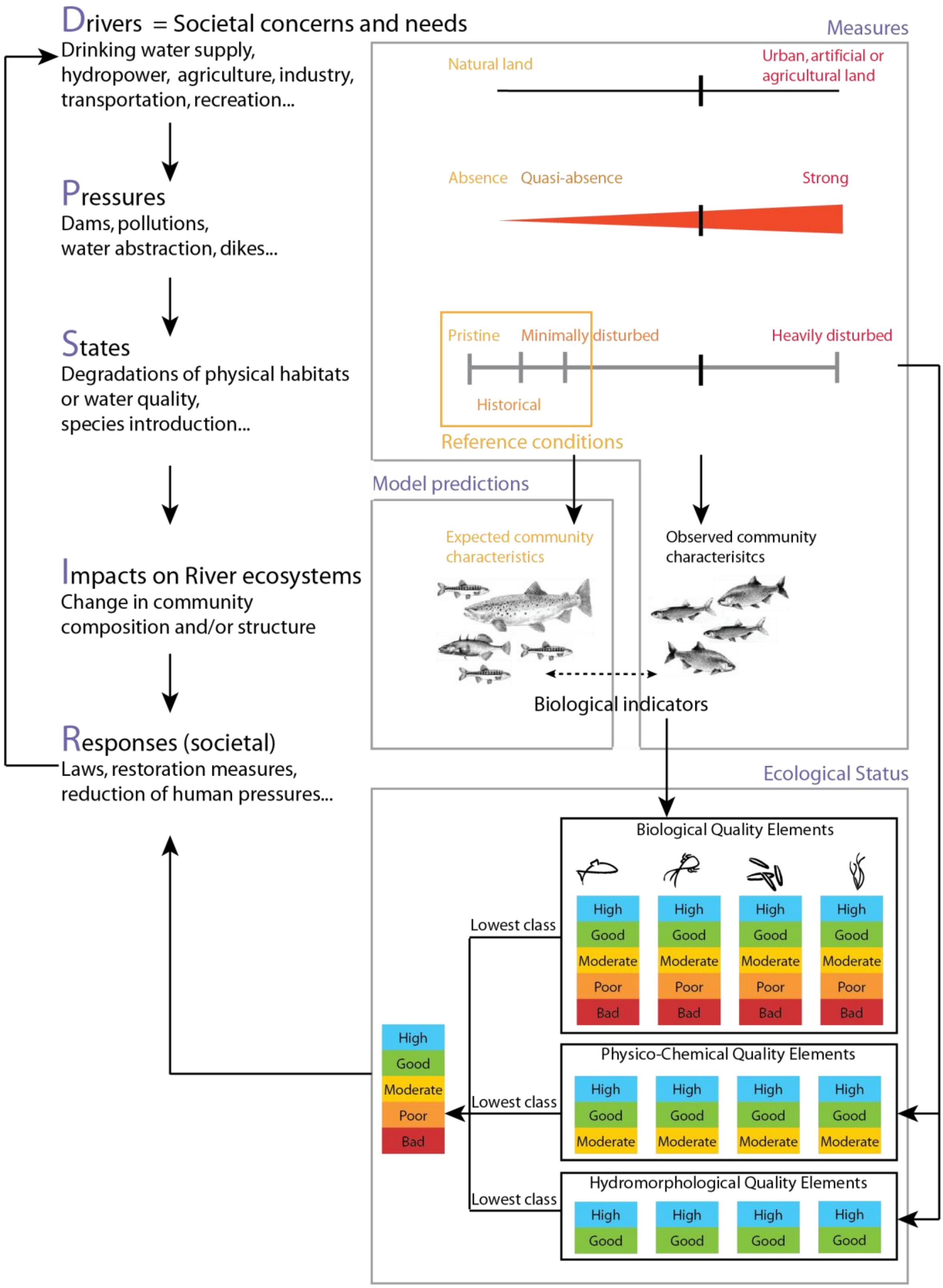


Figure 2. Bio-indication scheme and the WFD "one-out-all-out" rule. The quality element with the worst ecological quality is used to classify ecological status.

In more practical terms, the purpose of bio-assessment is to assist water management, improving the status of waters by identifying and quantifying existing impacts on aquatic biological communities and in some cases the likely causes of impacts (anthropogenic pressures) (CIS 2012). Bio-assessment is also used to check whether "good status" objectives are achieved (classification of the sites). Given the current European economic situation, bio-assessment provides decision-making tools (biological indicators) for water managers to facilitate prioritization and financial resource allocation for river management. Indeed, before deciding to implement measures to improve the ecological status of this or that river, it is critical for water managers to recognize the different situations and their potential chances of successful improvement. For instance, situations where the combination of several small human-induced disturbances have tremendous impacts on river communities could be distinguished from cases where a single substantial human-induced modification of the rivers has a relatively low impact on the river biota.

1.1.2 Underlying ecological concepts

"From a management perspective, bio-indicators inform our actions as to what is and is not biologically sustainable."
Holt & Miller 2011

Borrowed from the field of human medicine, excellent "health" of an ecosystem is conceptually defined as its "biotic integrity", i.e. "the ability to support and maintain a balanced, integrated, and adaptive community with a biological diversity, composition, and functional organization comparable to those of natural aquatic ecosystems in the region" (Karr et al. 1986). More commonly, it refers to ecosystem functioning within "normal" ranges, i.e. minimal deviation from the reference condition (Karr 1991; Karr & Chu 2000; Hamilton et al. 2010). Biological assessment (or bio-assessment) interprets the ecological health (also called quality) of a system from its resident biological communities (biota, e.g. fish, insects, algae, plants). River ecosystems are highly heterogeneous in space and time (Leopold et al. 1992; Gordon et al. 2004; Allan & Castillo 2007) and identifying the factors and processes that govern the structure of ecological communities (species composition, diversity and relative abundance of given species) is fundamental to assess river health. Therefore, bio-assessment is intimately linked to community ecology theory.

The "niche" theory suggests that species occur in a multi-dimensional space defined by both biotic and abiotic environmental factors where their traits (ecological and biological functions; Stearns 1992) allow them to survive and persist (Hutchinson 1957).

In river ecology, temporal scales range from days to centuries, while spatial scales range from microhabitat to catchment approaches. At the largest scales, historical community structure, regional species distributions (regional pool), phylogenetic history of species assemblages (Webb et al. 2002), species dispersal and colonization abilities (Rickliefs 1987), and long evolutionary events changing geography and climate (Rickliefs 2004) considerably constrain community structure.

At smaller scales, only species of the regional pool that present traits matching the habitat characteristics of the "environmental filter" can pass through (Fig. 3; Tonn et al. 1990; Poff 1997). As a result, spatio-temporal habitat heterogeneity is a framework in which evolution occurs selecting characteristic life history strategies of organisms and shaping the local community organization ("habitat templet theory"; Southwood 1977; Townsend & Hildrew 1994).

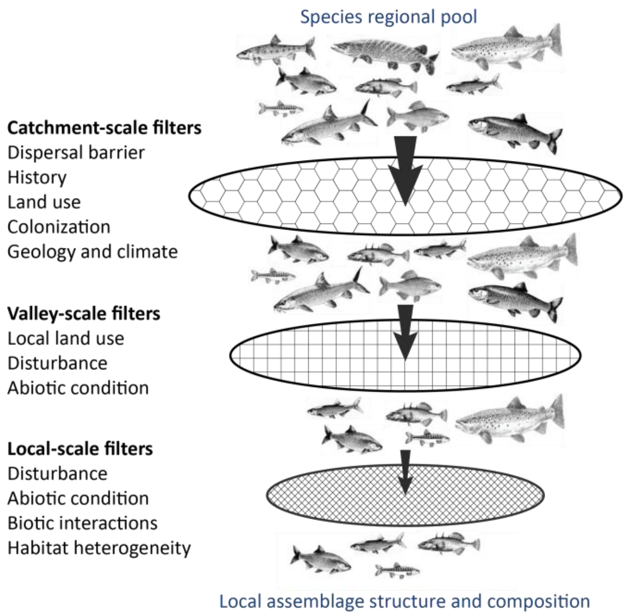


Figure 3. Hierarchical landscape filter theory. Adapted from Poff 1997 and Tonn et al. 1990.

Organisms are adapted to abiotic factors within a range of variation. As a result of human or natural disturbances, changes in abiotic factors outside the range of this variation may result in a shift in the distribution and abundance of organisms ("alternative stable states" theory; Suding et al. 2004). Either as a result of death or migration, certain species that are unable to tolerate substantial environmental changes could have their number of individuals decrease or disappear completely from the area concerned (McCormick et al. 2001; Wang et al. 2003; Pont et al. 2006; Pont et al. 2007). Therefore, the relative abundance and the number of certain species or groups of species are assumed to measure the degree of ecosystem integrity. However, natural disturbances such as drought or flood are sources of habitat diversity that also alter ecosystem processes and shape river communities (Sousa 1984). Consequently, human modifications regulating natural disturbances such as flow regimes could disturb the river dynamics and modify river communities (Poff et al. 1997).

1.1.3 Biological indicators

"Metrics measure something specific, while indicators are supposed to tell us something different from what they actually measure."
Daan 2005

A century-long story

Reflecting the close relationships between rivers and humans, biological monitoring appeared long before the WFD, and the use of aquatic communities to assess river health has a long worldwide history. It began in the second half of the nineteenth century to answer public health problems (e.g., cholera, dysentery, typhoid fever) and focused on micro-organisms and bacteriological aspects (Hynes 1960).

In the 1900s, in Germany and later in other European countries, public health and economic concerns stimulated a scientific research approach that used for the first time aquatic communities as indicators of river pollution related to urban and industrial discharges (including macrophytes, macroinvertebrates, and fish; Kolkwitz & Marsson 1902).

Since the 1970s, scientific and public environmental awareness as well as involvement in reestablishment of "ecosystem integrity" considerably increased (e.g. for the USA the Clean Water Act in 1972; the National Wildlife Refuge System Improvement Act in 1977) and the vision of human impacts on aquatic systems was enlarged, leading to considering other human disturbances than pollutions in bio-assessments (e.g. channelization, dams or deforestation). For instance, the Index of Biotic Integrity (IBI; Karr 1981) was the first multi-metric index to integrate several metrics based on fish functional guilds (e.g. trophy or tolerance to degradation) to describe fish communities' responses to human disturbances.

Since 2000, numerous biological indicators using the four BQEs have been developed to fulfill the new WFD requirements (Hering et al. 2004; Palmer et al. 2005; Furse et al. 2006; Schmutz et al. 2007). Numerous indices were derived from the IBI (Oberdorff & Hughes 1992; Hugueny et al. 1996; McCormick et al. 2001) and were grouped under the common designation: multi-metric index, such as the European Fish Index (Pont et al. 2006; Pont et al. 2007), the French fish-based index IPR (Indice Poisson Rivière; Oberdorff et al. 2002) and its recent improvement IPR+ (Pont et al. 2012).

Current needs and challenges for bio-indication

Bio-indicators should provide an unbiased assessment of human impacts on the ecological quality of rivers by (i) **observing the characteristics of biological communities of the river of interest**, (ii) **defining** how these river communities should look in the absence of human disturbances (i.e. **reference conditions**), (iii) developing and computing **metrics and indexes** to give a synthetic vision of the deviation of the observation from the reference condition, (iv) acknowledging the **uncertainty about this particular vision** in order (v) to facilitate decisions on the **status of the river** and the **measures needed to improve** its ecological integrity. Each of these steps determines the relevance of the assessment and of the resulting management measures. Although the bio-assessment field has been considerably advanced with the implementation of the WFD, many challenges and issues of debate remain (Hering et al. 2010).

What are the relevant indicators of the impacts of human disturbances on river ecosystems?

Historically, river biological integrity was examined through metrics focusing on "indicator organisms" (generally sensitive taxa, e.g. the Saprobic index and further developments; Kolkwitz & Marsson 1902; Pantle & Buck 1955; Zelinka & Marvan 1961; Sládeček 1965; Rolauffs et al. 2004), but nowadays two main approaches stand out. One describes the ecological integrity of the ecosystem through metrics based on the community taxonomic composition (e.g. relative abundance of the Plecoptera taxa), i.e. the **"compositional" approach** (Heino et al. 2005; Johnson & Hering 2009), while the other is based on the functional structure of communities (e.g. the number of insectivorous species), i.e. the **"functional" approach** (Usseglio-Polatera et al. 2000; Doledec et al. 2006). This second approach assumed that combinations of functional traits (ecological and biological) are selected by habitat conditions through the survival ability of individual organisms relative to others (Southwood 1977; Townsend & Hildrew 1994). **Multi-metric indexes** combined several compositional or functional metrics that reflect some aspects of the structure, function, or other characteristics of the aquatic communities. The consideration of several metrics is assumed to better evaluate the ecosystem condition than a single metric and to represent a wide range of community responses to human disturbances impacting the area concerned (Karr 1991; Karr & Chu 2000).

In addition, to detect the effects of multiple human stressors on aquatic ecosystems may require the use of multiple assemblages (Johnson & Hering 2009) and the one-out-all-out rule assumes that each group provides particular information on community integrity and responds specifically to human disturbances.

Consequently, a vast collection of biological monitoring approaches exists and the **selection of appropriate metrics and biological groups** to assess the ecological health of rivers is often the subject of considerable discussions (e.g. Bunn & Davies 2000).

Observing the local river biota, which human impacts are detectable?

Spatiotemporal heterogeneity of rivers shapes community structure through hierarchical filters (Fig. 3; Poff 1997). Biological indicators are assumed to be influenced by human disturbances and insensitive to other "natural" factors known to shape communities (e.g. Stoddard et al. 2006), e.g. spatial and temporal heterogeneity of the natural environment such as the river size or the hydrological regime. To acquire such biological indicators, the factors that govern the structure of ecological communities (species composition, diversity and relative abundance of given species) have been widely studied (e.g. Chessman et al. 2006; Schmutz et al. 2007; Leunda et al. 2012). However, the **relative influence of the natural environment and the different spatial scales of human pressures** on river communities have been rarely studied (e.g. Lammert and Allan 1999; Wang et al. 2003). **Complex interactions** among these different factors have been suggested (e.g. Snyder et al. 2003) but almost never quantified. In addition, since European rivers are frequently impacted by multiple stressors, **single effects of stressors on BQEs** have not been sufficiently assessed (Hughes et al. 2009). Nevertheless, such analyses will improve the understanding of pressure–impact relationships and provide useful information to predict and manage the impacts of human disturbances on river local communities.

How can we take into account the imperfections of our data, knowledge and methods?

Since biological indicators are always affected by spatiotemporal natural variability and human pressures (Stoddard et al. 2006), these two sources of community variability must be distinguished when assessing the impact of human disturbances (Sandin & Verdonschot 2006). To account for the natural heterogeneity of river communities, the WFD recommends using the RCA. Standardized sampling methods and processing protocols are applied to acquire comparable biological and abiotic data on national river sites (for instance, see AFNOR 2007 for diatoms and AFNOR 2004 for fish). A number of sites that are minimally exposed to human-induced stressors (reference sites) are sampled. The ecosystem integrity is defined as the range of biota that occurs at these sites (reference condition; Stoddard et al. 2006). Two types of method are mostly used:

- In "predictive multivariate" methods, the variability in the biota at reference sites is explained as a function of abiotic factors (e.g. distance to the source, geology or temperature) using statistical models (e.g. Hawkins et al. 2000; Oberdorff et al. 2002; Clarke et al. 2003; Pont et al. 2006). In contrast to reference sites, the test sites could be exposed to some degree to stressors that could affect the biota. The resulting models are then used to predict what the biology should look like at a test site if it has been unimpaired by human activity.

- The "typological" method groups water bodies according to their physical and morphological attributes (e.g. salinity, catchment size, altitude) and defines the range of reference biota for each type of river (e.g. Verdonschot & Nijboer 2004; Lorenz et al. 2004, Ferreol et al. 2005).

Deviation between the expected (reference condition) and actual communities represents the level of impacts caused by human degradations (Bailey et al. 1998). The Ecological Quality Ratio (EQR; European Union 2000) is then generally obtained dividing the observed metric by the reference value. These scores are easy to interpret as they give comparable measures of human stressor impacts on ecosystems.

The reference condition represents the baseline of the assessment. Since both typology and predictive models are a simplification of the actual situation and rely on sampling data, a part of the variability of the reference conditions (i.e. natural spatiotemporal variability of the river ecosystem) is not encompassed and could lead to uncertainty in bio-assessment.

More generally, "**biological indicators necessarily tend to simplify ecosystems into a single number** that is a broad indicator of some notion of quality" (Hatton-Ellis 2008). Consequently, **multiple sources of uncertainty** could have repercussions on bio-indicator scores (e.g. sampling design and effort, model predictions) and thus affect the final diagnostics and water-manager decisions (e.g. Clarke & Hering 2006). Therefore, acknowledgment of the uncertainty of the assessment methods is a major challenge for WFD implementation (Hering et al. 2010). Yet, few indicators have been evaluated regarding these uncertainties (e.g. Ostermiller & Hawkins 2004; Bady et al. 2005; Carstensen 2007).

What are the effects of improvement measures on river communities?

Following the scientific and public awareness of human-induced damage to Earth's ecosystems, "restoration ecology is likely to be one of the most important fields of the coming century" (Hobbs & Harris 2001). One major challenge is that communities often do not respond predictably to improvement measures and could lead to unexpected results (Suding et al. 2004). Once human-induced changes in abiotic factors have resulted in degradation of river community characteristics, it is not certain that the restoration of the former conditions will lead to the return of the reference biota. Therefore, **collecting and making accessible the outcomes of river restoration measures is essential for improving restoration understanding and designs** (Palmer et al. 2007).

1.2 OBJECTIVES STRUCTURING THIS THESIS

This thesis deals with several aspects within this very broad framework by integrating what we learned from previous biological indicators as well as suggested solutions and recommendations for future developments or uses of biological indicators. In this general context, three objectives were addressed:

(1) Test some of the assumptions that influence the choices of biological groups and community descriptors to assess the integrity of river ecosystems.

(2) Use the IPR+ index to develop a method to account for biological indicator uncertainty related to the prediction of reference conditions and address two aspects of the uncertainty in bio-assessment: the predictive uncertainty and the temporal variability of bio-indicators.

(3) Regarding restoration, through the example of weir removal, review what has been learned about the effects of improvement measures on the river community using a conceptual framework.

The results of this work are presented and discussed following these three main objectives.

In the **first section**, the relationships of river communities and related bio-indicators with natural and human factors are examined. Several assumptions that determined the scales, biological groups and the methods used to deal with bio-assessment of river ecosystems are tested.

The first objective is to compare and understand the extent to which human disturbances and natural environment factors are able to explain spatio-temporal variability in community composition (***P2***). Human disturbances are described at three spatial scales and the natural temporal variability of river communities is illustrated.

The second objective is to use a comparative approach to provide elements for the determination of relevant indicators of human disturbance impacts on river communities (***P1***).

Functional metrics, compositional metrics and indexes based on four biological groups are considered (macrophytes, macroinvertebrate, diatoms and fish). The impacts of the combination of multiple and different types of human pressures on river biota are also addressed.

In the **second section**, as part of the development and validation of the IPR+ index, two aspects of the uncertainty related to bio-assessment are examined. One is related to the prediction of reference conditions and the second is related to the natural temporal variability of the bio-indicator.

The first objective is to present and discuss the method developed to account for the uncertainty related to the prediction of reference conditions (***P3***).

The second objective is to examine the temporal variability and predictive uncertainty of the IPR+ score (***P3***, ***P4***) in response to human disturbances and environmental factors.

The **third section** concerns the effects of ecological health restoration and improvement measures on the river biota. A conceptual framework is used to review and structure the knowledge and observed changes in river biota after the removal of small dams and weirs (***P5***).

Finally, the last section discusses the findings of this thesis and their implications for the assessment of the ecological status of rivers and draws some perspectives for future research in this field.

2. COMMUNITIES, FUNCTIONS AND BIO-INDICATION

"Without the moss in the tundra, the cutthroat in the mountain stream, and the canary in the coal mine, we may not recognize the impact of our disturbances before it is too late to do anything to prevent them."
Holt & Miller, 2011

Knowledge and understanding of the links between river communities and their habitat increase the probabilities of successful bio-assessment, river management and ecological improvement measures (Bond & Lake 2003; Naiman & Latterell 2005; Chessman et al. 2006; Schmutz et al. 2007; Leunda et al. 2012). River communities are shaped by river habitats through hierarchical nested relationships that are highly variable in time and space (Towsend & Hildrew 1994; Poff 1997).

Correspondingly, the idea has emerged that human impacts on the local river biota can be comprehended at different spatial scales (e.g. reach, valley or catchment; Hamilton et al. 2010). Several authors advocated that since the spatial scale of stressors has increased, the spatial and temporal scales of assessment and management should be increased as well (Verdonschot 2000). Others consider that knowledge gained from research at finer spatiotemporal resolutions may be more valuable and useful for dealing with specific environmental management issues (Leunda et al. 2012). Anthropogenic impacts on river communities have been largely documented at the short time scale and local scale (Fausch et al. 2002; Durance et al. 2006) and are now well documented at large scales also (segment and catchment) (Bis et al. 2000; Hrodey et al. 2009).

To our knowledge, only a few studies have attempted to compare the ability to explain the spatial variability in river biological assemblages at these different scales (Lammert & Allan 1999; Moerke & Lamberti 2006; Nerbonne & Vondracek 2001; Richards et al. 1997; Sály et al. 2011; Wang et al. 2003). Most of these previous studies did not distinguish natural (noise) and human sources (signal) of community variability when assessing human impacts on ecosystems. Yet, for river management purposes, the latter are of prime interest as they represent meaningful triggers for stakeholders to improve or maintain ecological quality of rivers. In addition, complex interactions among these different factors have been suggested (e.g. Snyder et al. 2003) but rarely studied and quantified. Although commonly recognized as a great source of variability in river community structures, the temporal dimension was not the primary objective of this work and this section is more an illustration of this aspect than an in-depth examination.

As mentioned before, to detect the changes of community structure in response to human disturbances, numerous biological indicators (e.g. multi-metric, predictive) based on different biological groups and types of community descriptor (e.g. compositional or functional metrics) have been promoted (Bunn & Davies 2000). The power of these methods to detect human-induced impacts on river communities is certainly variable. However, responses of indicators to human disturbances in rivers have been frequently analyzed separately (e.g. Archaimbault et al., 2010; Besse-Lototskaya et al., 2011; Lacoul & Freedman, 2006; Yates & Bailey, 2010) but infrequently compared (Heino et al. 2005; Hering et al. 2006; Johnson et al. 2006a; 2006b; Johnson & Hering 2009; Hughes et al. 2009; Justus et al. 2010). Moreover, due to the rarity of large data sets, the common case of multi-impacted sites was generally not differentiated from individual pressure impacts on the community.

To address these aspects, this section presents and discusses the results of the studies conducted in articles ***P1*** and ***P2***. The objectives were to **(1)** study the **relative influences of human disturbances** measured at **three different spatial scales** (reach, riparian corridor, and catchment) and the **"natural" environment** on river biological community composition (macroinvertebrate and fish), **(2)** illustrate the **natural temporal variability of fish community composition** and **(3) compare the responses of a collection of biological indicators** (different biological groups and types of indicator) to human disturbances.

2.1 RIVER COMMUNITIES' RESPONSES TO SPATIAL HETEROGENEITY (***P2***)

To analyze the influence of "natural" heterogeneity (also called "physiography") and spatially different anthropogenic pressures (catchment, riparian corridor, local) in shaping river biological assemblages, data including fish and macroinvertebrate samples as well as environmental abiotic factors, local human disturbances and land use were compiled for 301 French sites (Fig. 1; ***P2*** p. 120).

Firstly, the ability to explain variations in **macroinvertebrate and fish assemblage composition** is compared between the three spatial scales of human pressures and the physiography. Secondly, after having accounted for the effect of physiography, the **relationships between anthropogenic pressures and assemblage composition** are analyzed. Finally, the **implications** of these results for river bio-assessment are discussed.

2.1.1 Relative influence of natural and human factors on community composition

The unique and shared influences of the three-scales of human pressures and physiography were analyzed using variance partitioning of the biological assemblage **composition**, i.e. taxa-relative abundances (Fig. 4; Peres-Neto et al. 2006; Borcard et al. 2011).

(1) It appears that in comparison to macroinvertebrate assemblage composition, **spatial variation in fish assemblage composition** is **better explained** by natural and human disturbance factors (as considered in this study; Fig. 1 and ***P2*** p. 120). This difference in explained variability might be due to the coarser level of taxa identification of macroinvertebrate assemblages (genus or family levels versus species level for fish). A part of the anthropogenic effects might be masked by pooling several species with different preferences and responses to pressures into a single family.

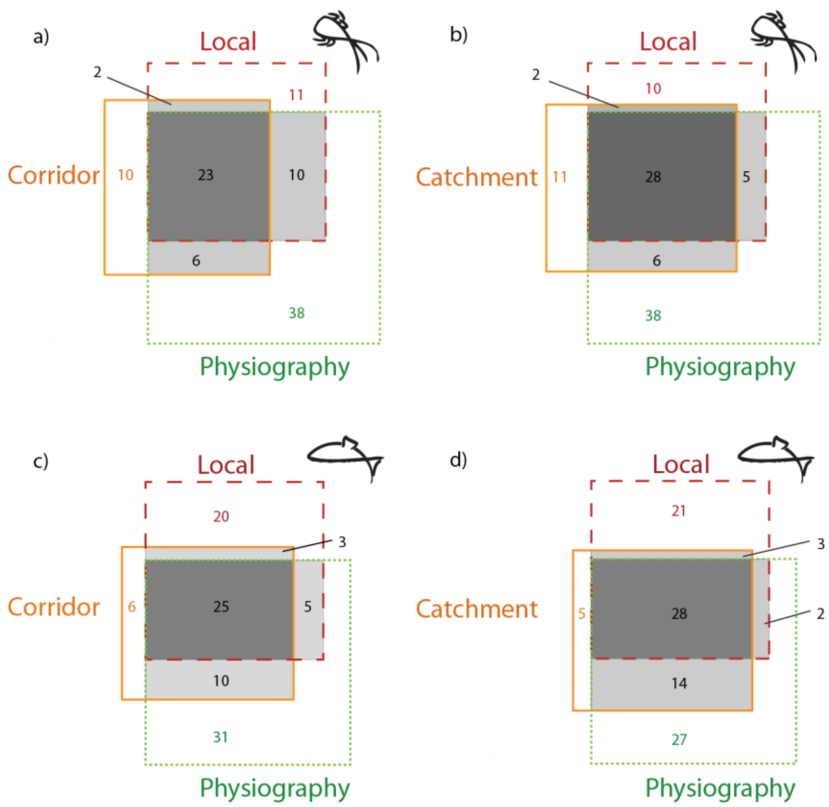


Figure 4. Venn diagrams representing the amount of variation in macroinvertebrate (a, b) and fish (c, d) composition explained by physiographic, anthropogenic factors (local, riparian corridor and catchment scales, see Fig. 1 caption p. 28 for pressure definitions), and their interactions. Each area is proportional to the share of inertia explained by the single factors (white area) or their interactions (grey areas). The numbers correspond to the percentage of the explained variation associated with each variable type. For both macroinvertebrate (fish) analyses, about 15% (30%) of the total variance was explained.

(2) As expected, for both BQEs, **a large part** of this explained share of inertia is described by **physiographic variables** (about 40% and 30% for macroinvertebrate and fish communities, respectively) (Fig. 4). This result is in agreement with previous studies showing that "natural" heterogeneity in the environment is a key parameter explaining spatial diversity of river community composition (e.g. Logez et al. 2010).

(3) Also, **complex interaction effects** among natural and human factors (about 40%) seem to play a major role in shaping communities (Fig. 4).

(**4**) The relative influences of anthropogenic pressures at the different spatial scales are different for macroinvertebrate and fish assemblages. Fish community composition appears to be more sensitive to local anthropogenic pressures (Fig. 4c, d), while land use variables and local scales variables have the same importance for macroinvertebrate community composition (Fig. 4a, b). Land-use variables reflect water quality degradations of the reach and upstream drainage basin area better than reach-scale hydro-morphological degradations (***P2*** p. 122). These differences are probably related to the fact that fish assemblages are globally more sensitive to hydrological and morphological degradations than macroinvertebrate assemblages (Hering et al. 2006; Justus et al. 2010). In addition, as previously mentioned, the response of macroinvertebrate assemblages to small-scale structure might be masked by the coarse taxonomic resolution.

2.1.2 Relationships between human pressures and community composition

The relationships between anthropogenic pressure factors and the spatial diversity of the two BQE assemblages have been described using partial redundancy analyses (Borcard et al. 2011, p. 171). This analysis displays the patterns of taxa composition uniquely explained by a linear model of the human pressure variables when the effect of the physiographic factors is held constant. For each BQE, one analysis per scale is conducted.

Community taxa composition ~ Human pressures | Physiography

As in the previous analyses and probably for the same reasons (mainly coarser taxonomic resolution), the share of the total inertia of community compositions explained by the analyses was lower for macroinvertebrates (13–14%) than for fish (23–28%).

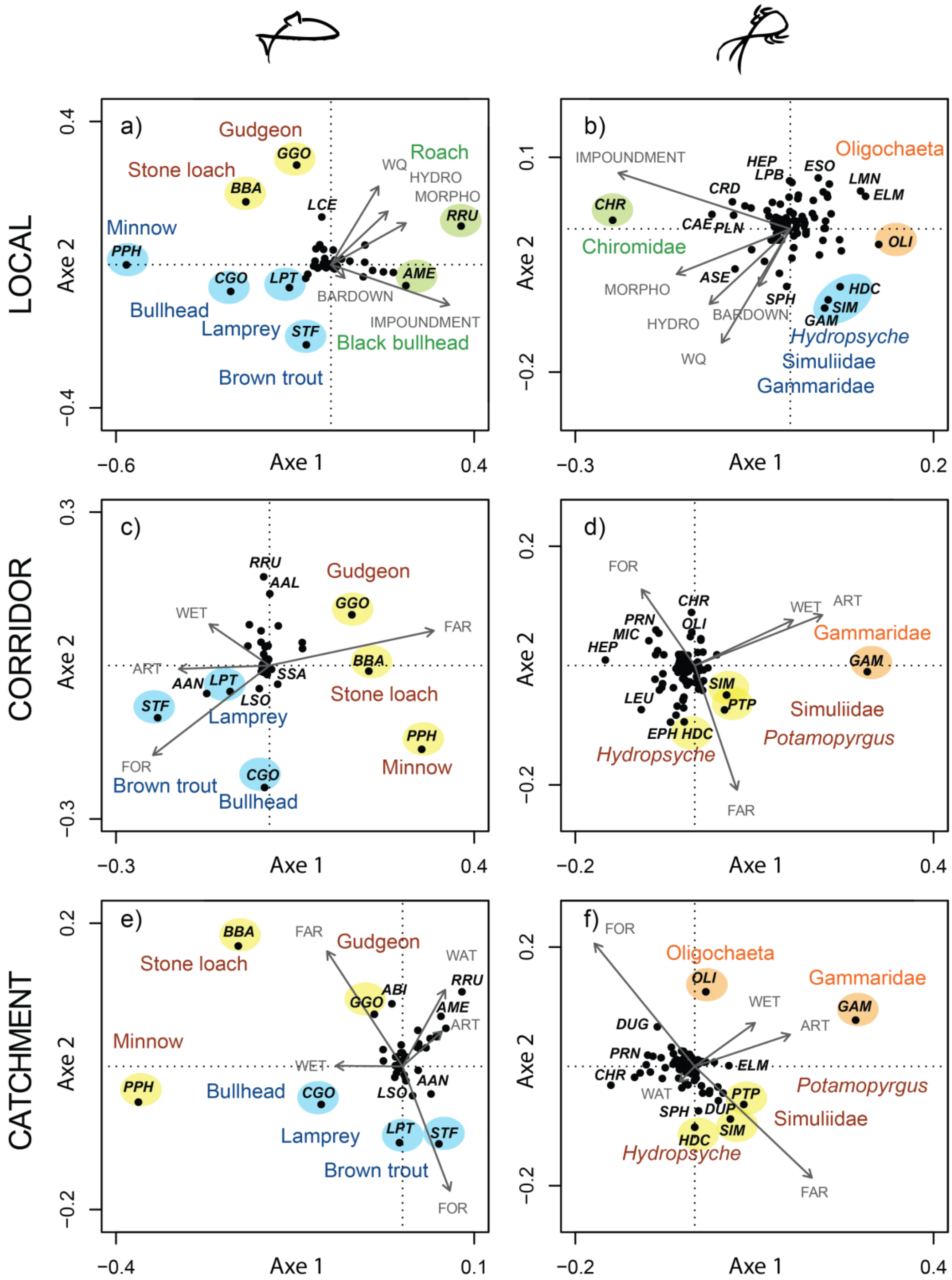


Figure 5. Biplots from the six independent partial redundancy analyses using anthropogenic variables at three spatial scales: local (a, b), riparian corridor (c, d) and catchment (e, f) (see Fig. 1 caption p. 28 for pressure definitions). Arrow length corresponds to the strength of the relationships among the variables and the axes. Three-letter codes represent fish species (ABI, *Alburnoides bipunctatus*; AAL, *Alburnus alburnus*; AME, *Ameiurus melas*; AAN, *Anguilla anguilla*; BBA, *Barbatula barbatula*; CGO, *Cottus gobio*; GGO, *Gobio gobio*; LPT, *Lampetra* sp.; PPH, *Phoxinus phoxinus*; RRU, *Rutilus rutilus*; SSA, *Salmo salar*; STF, *Salmo trutta fario*; LCE, *Squalius cephalus*; LSO, *Telestes souffia*) and macroinvertebrate taxa (ASE, *Asellidae* Gen. sp.; CAE, *Caenis* sp.; CHR, Chironomidae Gen. sp.; CRD, *Corixidae* Gen. sp.; DUG, *Dugesiidae* Gen. sp.; DUP, *Dupophilus* sp. Ad.; ELM, *Elmis* sp. Ad.; EPH, *Ephemera* sp.; ESO, *Esolus* sp. Ad.; GAM, *Gammaridae* Gen. sp.; HEP, *Heptageniidae* Gen. sp.; HDC, *Hydropsyche* sp.; LPB, *Leptophlebiidae* Gen. sp.; LEU, *Leuctridae* Gen. sp.; LMN, *Limnius* sp. Ad.; MIC, *Micrasema* sp.; OLI, *Oligochaeta* Gen. sp.; PLN, *Planorbidae* Gen. sp.; PTP, *Potamopyrgus* sp.; PRN, *Protonemura* sp.; SIM, *Simuliidae* Gen. sp.; SPH, *Sphaerium* sp.).

(1) However, **common patterns of response to pressures** were observed for fish and macroinvertebrate communities. Biological community distributions along the pressure gradients are mostly coherent with bio-ecological knowledge on fish and macroinvertebrate taxa (Fig. 5).

(2) The **presence of an impoundment** emerges as the main human pressure factor shaping the communities at the local scale, **followed by water quality** and morphological pressure gradients (Fig. 5a, b). The presence of an impoundment is known to have major impacts on river assemblages (Baxter 1977; Ward & Stanford 1979; Tiemann et al. 2004) as it considerably changes river functioning by shifting from free-flowing water to stagnant water systems.

(3) At **larger scales** (riparian corridor and catchment), fish and macroinvertebrate community composition appears to be greatly influenced by a **common gradient from forested cover to agricultural land use** (Fig. 5c, d, e, f). For instance, stenotherm-intolerant fish species (such as brown trout, bullhead and lamprey) preferring to spawn in running water were relatively more abundant in rivers surrounded by forested land cover, probably in relation to the dominance of small streams. More tolerant taxa, such as gudgeon, stone loach and minnow for fish and Simuliidae, Leuctridae, Hydropsyche, Potamopyrgus and Dupophilus genera for macroinvertebrates appear to be related to agricultural land cover.

In addition, an increase in corridor artificial and wetland cover appears to be another important gradient influencing macroinvertebrate assemblage composition (Fig. 5d), particularly the abundance of Gammaridae. Oligochaeta are negatively correlated to reach-scale pressures and farm land use and positively correlated to catchment forest cover and artificial land use. Oligochaeta are often recognized as species tolerant to water quality pollution and eutrophication (Lafont et al. 1996; Verdonschot 2006). Consequently, it appears likely that this relationship is explained by the presence in forest streams of more intolerant species such as Naididae (Tachet et al. 2006; Verdonschot 2006).

Box 1

Macrophytes and diatoms

For comparison, the same analyses were later conducted for macrophyte and diatom communities. The percentages of the variability explained were particularly low for the different analyses. The quality of the analyses might be improved by the addition of supplementary variables that are meaningful for these groups (for instance for diatoms, describing the water mineral composition, e.g. silica). However, the results were mostly comparable with macroinvertebrate and fish analyses, with at the local scale, predominant impacts of degradations of water quality and the presence of an impoundment on the assemblage composition and at broader scales, a gradient from artificial and farming area to forested land cover structuring the compositions of these communities.

2.1.3 Implications for bio-assessment

The findings of this study strengthen the idea that "natural" spatial heterogeneity should always be considered and recognized from human-induced spatial heterogeneity when looking at the impacts of human disturbances on river ecological quality, as it explains a large proportion of community spatial variability.

It appears that land uses and local pressures both significantly explain spatial variations of fish and macroinvertebrate community composition. Local variables are essential to describe responses of river ecosystems to human-induced pressures. Although catchment and riparian proportions of the forested area appear to be useful surrogates of the global water quality degradation of the upstream river, they should always be combined with information on local scale pressures.

Fish assemblages seem more influenced by reach-scale human pressures than land use, whereas macroinvertebrate assemblages present comparable sensitivity to land use and reach-scale variables. Given their different responses, the use of multiple assemblages may be appropriate to monitor river ecological quality.

Moreover, the existence of multiple interaction effects among physiographic factors and human pressures on river biological communities illustrates why it is so difficult to establish simple pressure–impact relationships for fish and macroinvertebrates, as pressure effects are generally difficult to separate. Consequently, in the common case of multi-impacted sites, it is very hard for water managers to determine the main pressure disturbing river ecological status. Future advanced research on complex interaction effects among human pressures and between pressures and physiography would advance both understanding and management of river ecosystems.

2.2 NATURAL TEMPORAL VARIABILITY IN RIVER COMMUNITIES

The inter-annual variability of fish assemblage composition in absence of changes in human disturbance was studied using the 10-year time series of fish, environment and human pressure data for 183 French sites (***P4*** data set; Fig. 1). The species composition matrix was transformed to acquire relative abundances and give low weights to rare species (Hellinger transformation Legendre & Gallagher 2001). The temporal dissimilarities in assemblage structure were quantified for each site as the average Morisita-Horn index among years because it is not influenced by species richness and sample size (Wolda 1981). Morisita-Horn dissimilarity indexes averaged 0.11, ranging from 0 to 0.33.

(1) Inter-annual variation of assemblage composition appears to **increase with human-disturbance intensity** (Fig. 6; $p<0.001$). This result is in accordance with previous studies assuming that human disturbance might destabilize river communities (Collier 2008; Franssen et al. 2011; Schaeffer et al. 2012). **(2)** Contrary to common assumptions (Horwitz 1978; Schlosser 1982; Oberdorff et al. 2001), fish community inter-annual variability was **higher downstream than upstream** in this study (Fig. 6; $p<0.001$). The instability of the assemblage might reflect the difficulty of achieving representative samples in larger rivers. Since the abundance and number of species are generally greater downstream (Huet 1954; Horwitz 1978; Oberdorff & Porcher 1992), the sampling effort might not always be sufficient to represent the whole community present in the river, resulting in increasing composition variability downstream.

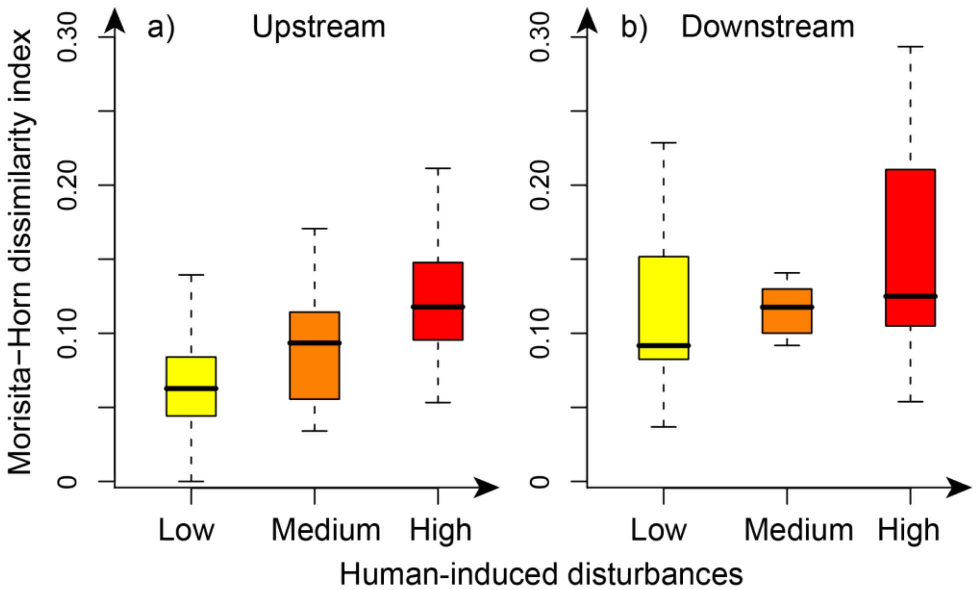


Figure 6. Species assemblage composition inter-annual variability relationships with human-induced disturbance intensity for upstream (a) and downstream sites (b).

In addition to inter-annual variability, seasonal variation in habitats highly influences fish communities as it is directly related to species' life cycles such as growth or reproductive period (Schlosser 1982). Moreover, depending on their life stages, fish may require different habitats. The impacts of these two levels of variability have been widely studied and are generally minimized in bio-assessment by using samples from the same season and year.

Implications for bio-assessment

The confusion of the effects of natural temporal variability and human impacts on the community structure could bias assessment. Natural temporal heterogeneity could be the result of multiple confounding sources such as exceptional climatic or hydrological events (e.g. flood, drought), varying sampling effort and stochastic biological processes (e.g. recruitment, competition). Therefore, such sources of variability of community structures should be considered when defining the expected biota in reference conditions and when developing bio-indicators to minimize the confounding effects (Mazor et al. 2009; Linke et al. 1999). This variability seems to be particularly important for downstream sites that are often subject to strong human-induced degradations and high sampling variability. As a result, bio-indicators and bio-assessment relying on fish communities are likely to vary more over time for downstream impacted sites than for upstream, more pristine sites.

Finally, in a context of global changes, this reflection also strengthens the need to sample reference sites through time to account for reference condition variation for future bio-assessments (Logez & Pont 2012).

2.3 DESCRIPTORS TO ASSESS RESPONSES OF RIVER COMMUNITIES TO HUMAN DISTURBANCES (***P1***)

The responses to human disturbances of 93 metrics based on macroinvertebrate, diatom, macrophyte and fish assemblages were compared (***P1*** pp. 108–109). The metrics tested comprised indexes described in the literature and compositional and functional metrics. Biological samples as well as information on the physiography and local human pressures impacting rivers were compiled for 290 French sites (Fig. 1; ***P1*** p. 107). In contrast to previous studies (Heino et al. 2005; Hering et al. 2006; Johnson et al. 2006a; 2006b; Johnson & Hering 2009; Hughes et al. 2009; Justus et al. 2010), natural variations in stream ecosystem functioning were differentiated beforehand from human pressure effects by standardizing the metrics (Oberdorff et al. 2001; Pont et al. 2006; 2007). All the results are presented for metrics and indexes standardized using this method (see ***P1*** pp. 108–109 for details). Therefore, the values of the subsequent metrics were assumed to vary according to the intensity of human pressures and independently of “natural” heterogeneity in the environment.

Three criteria were used to describe the responses of the metrics and indexes to human pressure spatial variations:

- a **significant difference** in metric values between weakly disturbed sites and sites strongly disturbed by the pressures considered.
- the **intensity of the responses** was measured as the discriminatory efficiency (DE; Ofenböck et al. 2004), i.e. the percentage of severely disturbed sites with metric values below (above) the 5^{th} (95^{th}) percentile of the slightly disturbed sites for increasing (decreasing) metrics.
- the **sensitivity** of the metric was defined as the lowest level of pressure detected by the metric (i.e. metric values significantly different from slightly disturbed sites).

Kruskal-Wallis non-parametric post-hoc tests were used to determine the effect of single and combined human pressures on BQE metrics.

First, the responses of the metrics of different "natures" are compared (functional, compositional or indexes). Second, the responses of the metrics are compared depending on the biological community they describe. Finally, the implications of these results for bioassessment are discussed.

2.3.1 Which type of metric is more appropriate to detecting human pressure impacts?

The 93 bio-indicators were based on bio-ecological functional traits or taxonomic composition or corresponded to published indexes. Sixty-six out of the 93 transformed metrics detected at least one of the anthropogenic pressures considered. All the indexes, two-thirds of the functional metrics and only one-third of the compositional metrics responded to a global degradation gradient (see examples in Fig. 7).

Indexes and functional metrics showed the **strongest responses** to human-induced degradation, whereas most of the compositional metrics showed a weak response to human pressures. **Functional metrics** seem to be generally **more sensitive** to human pressures (i.e. detect weaker disturbances) than indexes and compositional metrics. This is in agreement with several authors suggesting that ecological and biological functional traits are well-adapted for large-scale approaches (Statzner et al. 2001) and are able to integrate more general phenomena than compositional metrics (Dolédec et al. 2006).

In addition, most of the **indexes** considered are significantly **affected by all the pressure gradients**. These results support the use of ecological and biological functional trait metrics to build multi-metric indexes to assess river biotic integrity.

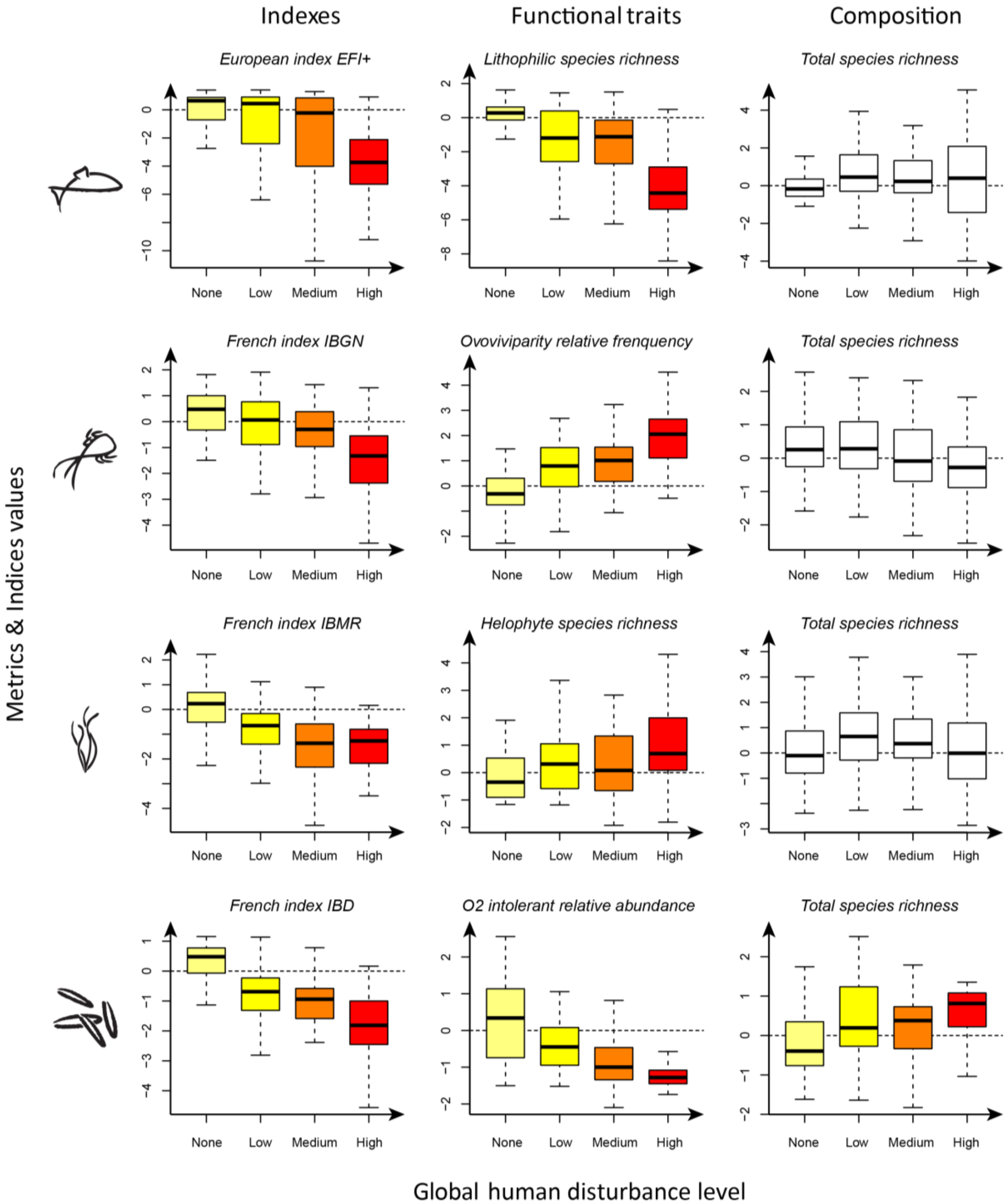


Figure 7. Examples of metrics and multi-metric indicator responses to global human-induced perturbations. Black and white plots (respectively, coloured plots) for nonsignificant (significant) responses. Pressure gradients represented by four classes from minimally disturbed sites (None) to highly impacted sites (High).

2.3.2 Do BQEs' metrics detect anthropogenic degradations in a similar way?

A common pressure hierarchy stands out for the four BQEs (Fig. 8). In agreement with previous studies, metric responses in terms of intensity and sensitivity were **stronger overall for global degradation than for specific pressures** (Hering et al. 2006) and among specific pressures, **water quality degradations** resulted in the strongest responses (Hering et al. 2006; Johnson et al. 2006b; 2009; Justus et al. 2010).

However, sensitivity and intensity of metric responses to human pressures **fluctuated among biological groups** (Fig. 8). **Diatom and macroinvertebrate metrics** appear to be **more sensitive** to the degradation of the overall condition of the river than fish metrics and reacted to lower levels of pressure. Macrophyte metric responses to global degradations were weaker and less sensitive than for the other groups. **Fish metrics** had the **strongest intensities**. These differences may be partly related to the migratory capacities of fish and their longer life cycles. Consequently, as long as favourable habitats and conditions are accessible for fish, changes remain undetected by metrics. When they are no longer accessible, fish metrics show dramatic responses resulting in strong responses to strong perturbations. Conversely, short-life-cycle and sedentary organisms such as diatoms are impacted by a lower level of pressure as soon as local conditions are degraded. The less sensitive but more intense responses of fish metrics would be better adapted to detecting high modifications, while diatom and macroinvertebrate metrics would be more useful in detecting the first impacts of degradations.

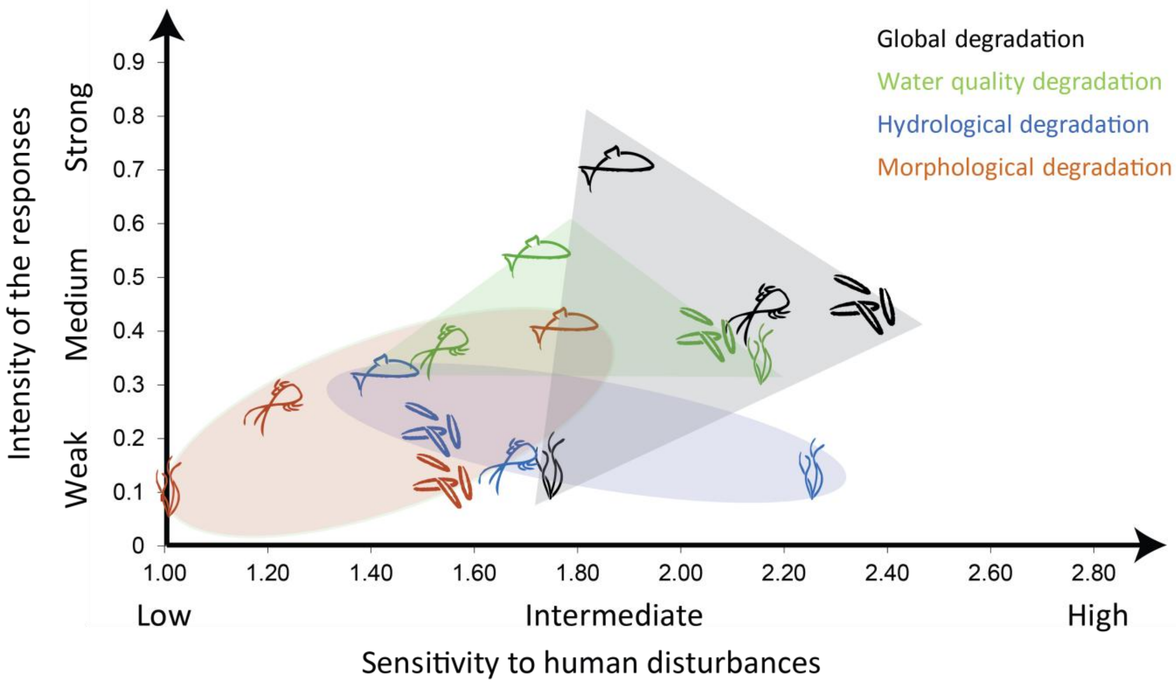


Figure 8. Median responses of tested metrics and indicators based on macroinvertebrate, macrophyte, diatom and fish river communities. Responses to global disturbances and three types of human-induced degradations (to global, water quality, hydrological and morphological) were described in terms of **sensitivity** and **intensity**.

The same kinds of patterns were observed for the other pressures (Fig. 8). **Diatom and macrophyte metrics** appear to be **more sensitive to water quality** (response to low to moderate levels of water quality degradation) than fish and macroinvertebrate metrics.

However, as in Johnson & Hering (2009), the strongest responses were observed for fish metrics.

Although responses were rather weak, contrary to Hughes et al. (2009), the present results suggest that **fish metrics** show the strongest responses to **hydrological perturbations,** while **macrophyte metrics** were the most sensitive. These differences may be explained by the authors' choice of variables describing pressures and the analysis settings. **Fishes** are the most impacted by **morphological degradations.** In addition, the results of this study confirmed the particular ecological impact of impoundments with relatively strong responses of the four BQEs to this pressure (***P1*** p. 111).

2.3.3 Are BQEs' responses the same when only sites impacted by one type of pressures are considered?

For hydrological and water quality degradations, the same analysis was carried out after removing the sites that were strongly impaired by other types of pressure (Fig. 9). As streams are frequently impacted by multiple stressors, the number of sites was too low to support statistical tests and single effects of morphological degradations were not tested.

For both water quality and hydrological degradations, the **numbers of significant responses decreased sharply** from pressure types to single pressure analysis and better responses in terms of intensity and sensitivity were still observed for water quality than for hydrological perturbations.

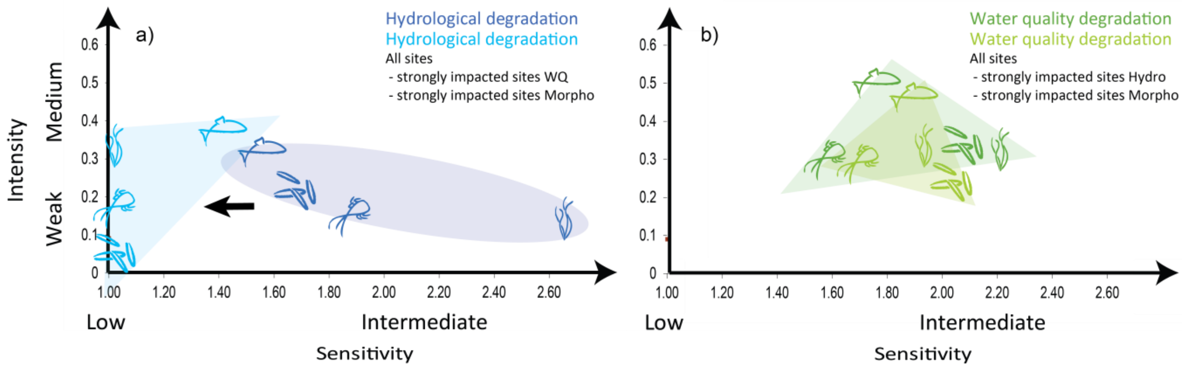


Figure 9. Median responses of metrics tested and indicators based on macroinvertebrate, macrophyte, diatom and fish river communities. Responses to hydrological (a) and water quality degradations (b) were described in terms of **sensitivity** and **intensity**. Dark colours for analyses including the 290 sites. Light colours for analyses including sites strongly disturbed by only one type of pressure.

Whereas the **intensity and sensitivity of the four BQE responses to water quality degradation were nearly unchanged** when removing sites strongly impacted by hydrological or morphological degradations, except for fish metrics, BQE metrics were less sensitive and showed weaker responses to hydrological degradations.

These results suggest that the effects on BQE metrics were mainly due to water quality degradation and not to a combined effect with hydro-morphological degradation in the first analysis. By contrast, for hydrological degradation, macrophyte, macroinvertebrate and diatom metric responses probably largely resulted from the impact of associated water quality and/or morphological degradation. These findings raise new issues about the relation between pressures. Unfortunately, the understanding of the combined or cumulated effect of several types of pressure on river aquatic assemblages is typically poor (Pont et al. 2007) and questions such as the existence of cumulative or multiplicative (i.e. interaction) effects remain unanswered.

2.3.4 Implications for bio-assessment

This study shows that intensity and sensitivity of the responses to human disturbance vary considerably among metrics. The selection of the metric to monitor the effects of stressors should not only focus on the BQE, but also on the nature of the metric (i.e. underlying processes, types and units). Indexes and functional trait-based metrics tend to detect human-induced changes better (stronger responses to a lower level of pressure) than compositional metrics. However, to expand this analysis, knowledge of biological and ecological traits needs to be improved, in particular for macrophytes.

This study shows that the four BQE metrics are impacted differently by pressures, even if the responses vary from one metric to another within a given group. This result reinforces the idea that different biological groups contribute new information on the nature and intensity of the human disturbance impacts on river communities and therefore should be considered jointly in river bio-assessment.

More generally, global and water quality degradations of the river appear to have stronger impacts on BQEs than morphological degradations and hydrological degradations. For multi-impacted sites, measures aiming to improve river hydro-morphological attributes might not show the expected effects if the water quality is still degraded.

Finally, given the present results, hydrological degradation effects will likely be confounded with water quality and morphological degradation effects on the biota if multi-impacted sites are not removed from the analysis. More generally, as river assessment research is turning towards multi-metric tools, it is of prime importance to be able to answer the following question before including metrics in indexes: Do the different types of pressure have additive, multiplicative or opposite effects?

Furthermore, this study has analyzed the influence of physiography on metrics for undisturbed conditions, but we believe that complex interactions exist between human pressure effects and the environmental diversity, i.e. the responses of aquatic assemblages to human pressure will be different depending on the physiography. Such interaction effects on BQE responses have not been sufficiently investigated to date and are needed to assess the ecological impacts towards restoration of water bodies.

2.4 CONCLUSIONS

(1) Natural spatiotemporal variability of abiotic features plays a major role in structuring river biological communities. Consequently, it could be confused with the impact of disturbances induced by human activities.

(2) Human disturbances act hierarchically on river communities. Although broader spatial scales encompass the global water quality of the upstream river, local variables are indispensable to describe impacts of hydro-morphological and local water quality degradation on river ecosystems. A hierarchy of disturbance types common to the different biological compartments stands out in our analyses.

(3) BQEs are impacted differently (response intensity and sensitivity) by human disturbances (spatial scales and types), supporting the notion that BQEs contribute complementary information on the level of degradation of river ecosystems.

(4) River communities result from **complex combined effects** between natural and human factors on the one hand and different types of human factors on the other.

(5) Biological metrics and indexes show highly variable responses (intensity and sensitivity) to human disturbance.

Human disturbance impacts on the river community appear to be the result of complex effects that are difficult to separate to provide a simple pressure–impact diagnosis. The biological communities and related metrics are globally more sensitive to a combination of different types of river degradation than to specific degradations. The degradation of water quality and the presence of an impoundment are the pressures with the most severe impacts. However, metric responses to human pressures vary among BQEs, suggesting that monitoring different BQEs enlarges the window through which we look into stream health.

Among the different possibilities tested, functional metrics combined into multi-metric indexes appear to be the most relevant to measure river biotic integrity.

Finally, to identify meaningful triggers for river quality improvement and management, natural environmental variability should be considered when defining the expected biota in reference conditions. The predictive method applied (Oberdorff et al. 2001; Pont et al. 2006; 2007) appears to be appropriate to achieve biological indexes and metrics independent of natural spatial variability while considering rivers as a continuum (Vannote et al. 1980).

3. UNCERTAINTY AND VARIABILITY ASSOCIATED WITH BIO-ASSESSMENT: THE IPR+ INDEX

“Counting fish is like counting trees…except they are invisible and they move.”
J.T. Schnute

Diverse sources of river ecosystem variability act together at different spatiotemporal scales and should be considered in river bio-assessment (***P1***, ***P2***). Today’s knowledge and understanding of the functioning of these complex systems are imperfect. Indicators aim at giving a synthetic view of this complexity to facilitate river management. Each step of the development of a bio-indicator (from sampling protocol to metric selection and index calculation) implies making choices that will influence the degree to which it reflects the actual complex situation (Clarke et al. 1996; Clarke & Hering 2006; Hering et al. 2010). Since they could bias the final assessment, the recognition of a lack of knowledge and uncertainty is indispensable to decision-making (Funtowicz & Ravetz 1993).

The sources of uncertainty in environmental management have been acknowledged as statistical uncertainty, subjective judgement, systematic error, incomplete knowledge, temporal variation and inherent stochasticity (Todd & Burgman 1998). In water body health assessment, the main sources of uncertainty are the incomplete acknowledgement of the **natural temporal and spatial variation of communities** and **errors** associated with **sampling processing methods** and **modelling (**Clarke & Hering 2006).

This part of the thesis was in line with the development and evaluation of the new fish-based index for French rivers, IPR+ (Pont et al. 2012). The need to recognize uncertainty in index validation raised two questions.

- **For a test site, what is the level of confidence of the “index score” and the subsequent “ecological status class”?**

This question could be reworded as “how large is the uncertainty around the index score?” Few studies have measured this uncertainty. Most have focused on errors related to monitoring and sampling strategy (Clarke et al. 1996; Ostermiller and Hawkins 2004; Bady et al. 2005; Staniszewski et al. 2006). For instance, Tomanova et al. (in revision) addressed the sampling error of fish community in large rivers using the French sampling method.

The use of statistical models to predict the expected biota in quasi-absence of human disturbances is one of the most widely employed techniques to discard the natural part of river community heterogeneity to highlight the effects of anthropogenic pressures on river ecosystems. For this type of predictive index, it appears essential to evaluate the often neglected uncertainty related to predictions of reference conditions (hereafter called "predictive uncertainty"). One part of model predictive uncertainty is irreducible, the uncertainty "in essence" due to the fact that a model predicts an average response while the underlying process is stochastic (i.e. a model simplifies reality). The other part depends on the quantity and quality of the data used to calibrate the model, the uncertainty "of ignorance" due to the fact that model parameters are estimated using a limited number of observations.

- **Is this evaluation consistent over time?**

In absence of human-induced changes, the evaluation of the ecological status should be consistent between years. Otherwise, in a case where human pressures are reduced or increased from one year to another, the effect of improvement or degradation of the ecosystem might be masked or confused with the "natural" temporal variability of the index. As illustrated in section 2.2, in addition to variation in sampling effort between years (Gotelli & Colwell 2001), the temporal variability of indicators may encompass the natural dynamics of aquatic communities (Bunn & Davies 2000; Ricklefs & Schluter 1993) and the effects of climatic hazards (Schaeffer et al. 2012). To our knowledge, the particular effect of temporal variability on bio-indicators has scarcely been studied (Collier 2008). However, some authors implicitly recognized and considered the effect of temporal variability within the process of metric selection (e.g. Hughes et al. 2004).

In this section **(1-2),** a **methodology to evaluate the uncertainty of a predictive multi-metric index** related to the **prediction of the metrics in reference condition** is described and discussed through the example of the IPR+ index (***P3***).

(3) The IPR+ index is evaluated regarding its capacity to detect impacts of different types of human disturbances on local fish communities, regardless of the natural environmental conditions (***P3***).

(4) In order to acknowledge the variability of the IPR+ index unrelated to human degradation and provide recommendations for potential users, **the inter-annual variability of the IPR+** and its **predictive uncertainty** were quantified and their **relationships with human pressure levels** and **natural environmental conditions** were studied (***P3, P4***).

3.1 THE IPR+ INDEX (*P3*)

The context of this study was the validation of a bio-indicator developed for the evaluation of French river ecological health using fish communities (IPR+). The first two stages of IPR+ development (Fig. 10) were conducted prior to this thesis (Pont et al. 2012). They are the basis of this work. The main principle and originalities of the IPR+ are exposed succinctly (see ***P3*** for a more detailed presentation of the index).

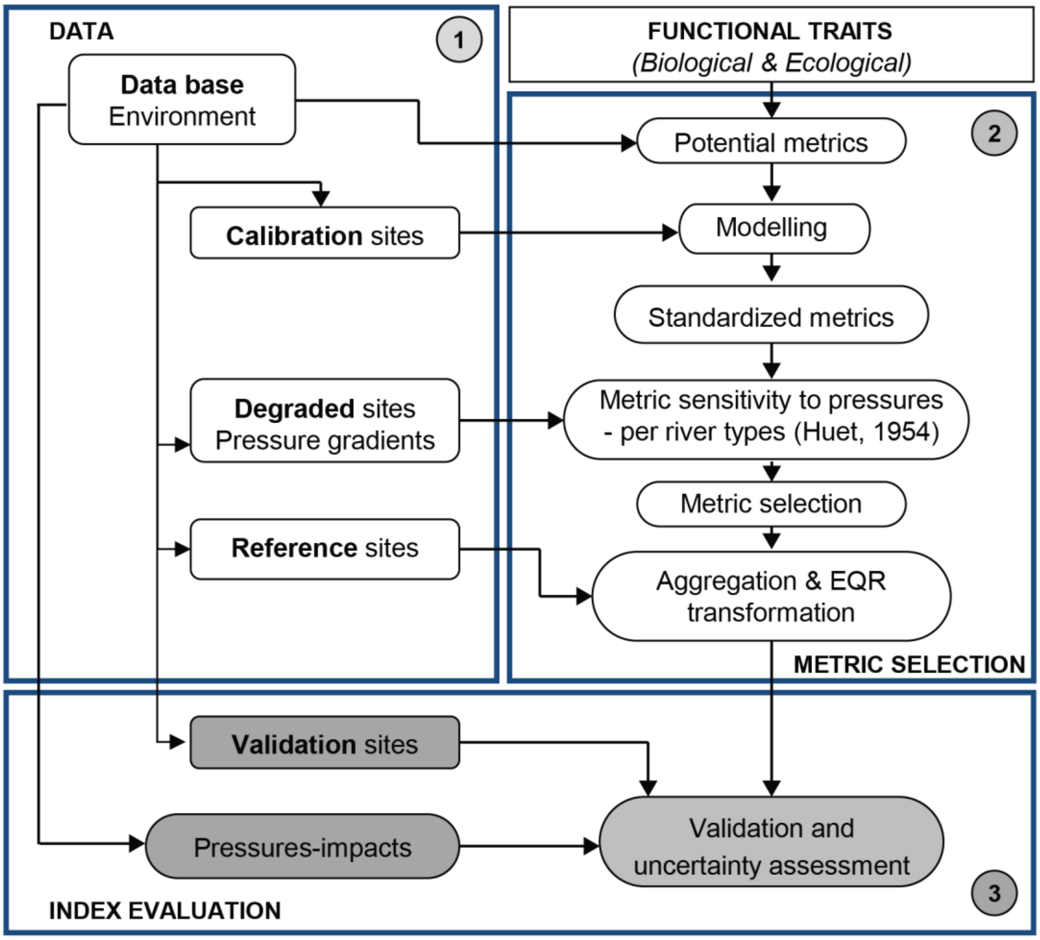


Figure 10. IPR+ development and validation. White tasks, done prior to this thesis. Grey tasks, done within this thesis.

The IPR+ is a multi-metric index based on fish functional traits. It was developed based on a national database (1654 sites, (Fig. 11) including fish samples (73 species represented in all), as well as information on six environmental features (e.g. geology or the temperatures and catchment precipitations of the 10 years preceding the sampling) and reach-scale human-induced degradations (Fig. 1; ***P3*** p. 149). Reference condition sites were defined from objective criteria based on pressure levels (Fig. 11).

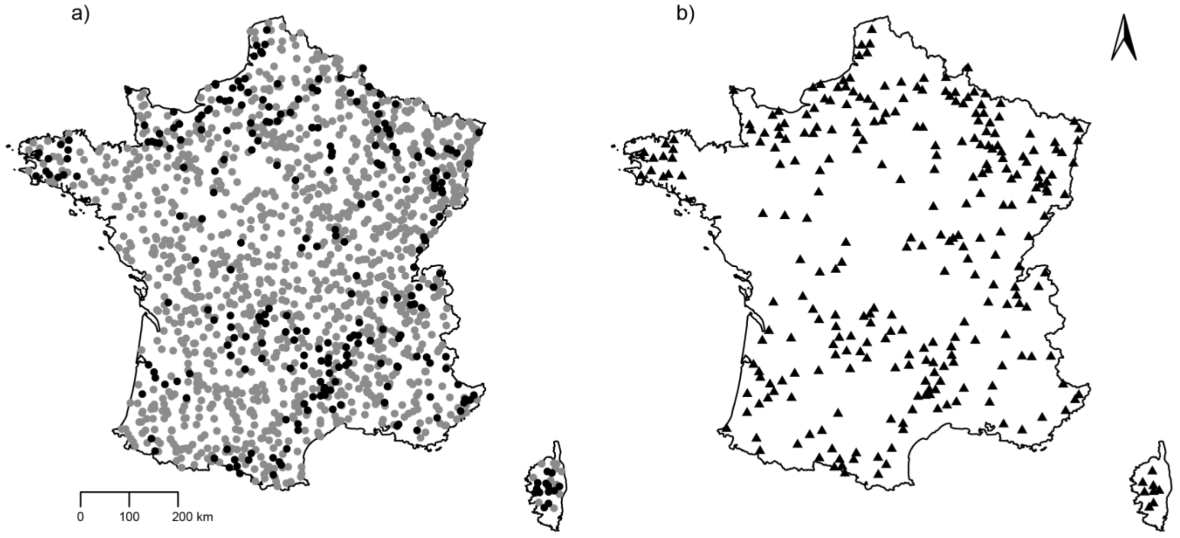


Figure 11. Localisation of the 1654 sites used for IPR+ development. (a) black dots, 266 reference sites and grey dots the other 1376 sites. (b) 278 model calibration sites.

Natural variability of the metrics is controlled to ensure that metrics measure the effect of anthropogenic disturbances while explicitly considering the river as a continuum (Vannote et al. 1980). Statistical models (generalized linear model; McCullagh & Nelder 1989), including the six environmental variables known to influence fish assemblages (Pont et al. 2005; Logez et al. 2012) and assumed to be unmodified or slightly modified by local anthropogenic activities, are used to predict the expected values of metrics in quasi-absence of human disturbance. Since the models include 10-year climatic variables, mid-term climate conditions are reflected and long-term climate change and its future impacts on reference conditions could be considered if necessary (Logez & Pont 2012).

Since French lowland rivers are rarely unimpaired, moderately impacted lowland sites are included in the model calibration data set to cover the largest gradient of environmental conditions. This compromise certainly decreases the power of the index by removing a part of the human disturbance effect from the signal, but justifies applying the index to lowland rivers (Fig. 11).

Functional metrics have been shown to make comparable rivers and sites having similar ecosystem functioning but different species pools (Lamouroux et al. 2002; Hoeinghaus et al. 2007) and to be better indicators of human-induced perturbation impacts on the biota than taxonomic metrics (***P1***). Accordingly, 11 metrics based on fish species biological and ecological traits (see ***P3*** for trait definitions p. 156) were selected regarding model quality and metric sensitivity to the different types of human disturbances. The 11 metrics entering IPR+ calculation are expressed in relative abundance (%N), richness (S) and relative richness (%S):

- trout juveniles (N%-Trout)
- oxyphilous species (N%-O2INTOL and S%-O2INTOL)
- species intolerant to habitat degradation (N%-HINTOL)
- species preferring to spawn in running waters (N%-RHPAR)
- species preferring to spawn in stagnant waters (S-LIPAR)
- tolerant species (S-TOL)
- stenothermal species (S-STTHER)
- omnivorous species (S-OMNI)
- intolerant species (S%-INTOL)
- limnophilic species (S%-LIMNO).

Each metric is standardized,

$$\textit{Standardized metric value} = \log(\textit{observed value} + 1) - \log(\textit{predicted value} + 1)$$

and divided by the median value of the reference condition sites and rescaled between 0 and 1 to obtain values expressed as EQR. To maximize the sensitivity of the final index to human-induced disturbances, the six metrics showing the lowest EQR values were retained for each site: two metrics based on abundance and four based on richness. The mean of these two groups of metrics was computed and then averaged to obtain the final index score.

For management purposes, the thresholds of the five ecological status classes were defined in agreement with European inter-calibration rules (Working Group Ecostat 2009; Willby et al. 2010).

3.2 A METHOD TO EVALUATE THE PREDICTIVE UNCERTAINTY OF A MULTI-METRIC INDEX

3.2.1 Why use a Bayesian framework?

The considerable advances in statistical theory and computing technology from the early 1980s onwards facilitated the development of Bayesian statistical methods and their application to relatively complex problems in environmental management (Brooks 2003). The main reason for using a Bayesian approach to build biological indicators is that it is particularly well suited for decision-making. In river management, decisions will need to be made before complete knowledge and perfect data are available (van der Sluijs 2007). Therefore, any decision relies on uncertain facts and this uncertainty quantification is crucial for evidence-based decision-making.

"The riskiness of a situation is fundamentally about the spread of a probability distribution." (Shaw & Woodward 2010). Bayesian and frequentist statistics principally differed in the definition of probability. Frequentist statistics define probability as the long-run expected frequency of occurrence (Hacking 1965). In a different way, the Bayesian view of probability measures the plausibility of an event given incomplete knowledge, i.e. a degree of belief (Finetti 1970). Consequently, most frequentist methods based on maximum likelihood or least-squares estimation involve setting the parameter values. By contrast, Bayesian methods explicitly use probability to quantify uncertainty resulting from imperfect knowledge, imperfect data, and environmental variability in the models (Gelman et al. 2004). Parameters are considered as random variables that are predicted in the form of probability distributions. This facilitates representing and taking into account the uncertainties related to models and parameter values. Posterior probability distributions for unknown parameters are explicitly derived from observed data and knowledge or belief through Bayes' theorem (Box 2).

Bayesian philosophy is based on the idea that more than what is contained in the data may be known. The underlying idea is that the prior (previous knowledge or beliefs on the probability distribution of the parameters) and the experimental results are both different views of reality. However, Bayesian methods are often criticized because of the subjectivity of the definition of prior probability, and for this reason generally uninformative priors are used to limit the level of subjectivity in the analysis (e.g. Punt & Hilborn 1998).

For comparison, a frequentist method to quantify the uncertainty of prediction for generalized linear models was developed using a Monte Carlo approach[3] (Manly 1997). For both methods, the calculation time was high and requires powerful computers. In the end, the results were quite similar. If both methods can quantify predictive uncertainty, why choose the Bayesian approach over the frequentist approach?

[3] Method developed by Maxime Logez (Irstea)

Within the Bayesian framework, powerful tools and software exist and are commonly used to deal with uncertainty issues (Spiegelhalter et al. 2003; McCarthy 2007). Consequently, the implementation of Bayesian methods did not necessitate the development of new statistical tools and was more accessible. In addition, this work is a first step in the development of Bayesian methods to account for predictive uncertainty in river bio-assessment. A further step will be to reduce IPR+ score uncertainty. Setting informative priors using knowledge on the relationships of metrics and environmental variables acquired from different data or using more sophisticated methods (i.e. hierarchical models; Gelman et al. 2004) might help to reduce the predictive uncertainty.

Box 2

Bayes' rule for distributions

For two events A and B, having probabilities P(A) and P(B) ≠ 0, respectively, the conditional probability of A given B is

$$\mathbf{P(A|B) = P(A \cap B) / P(B)}$$

where $P(A \cap B)$ is the probability that A and B events both occur and P(A|B) is the chance that event A will occur, given the fact that B has already occurred. The same relationship can be written for $P(B|A) = P(A \cap B) / P(A)$ leading to **Bayes' rule**:

$$\mathbf{P(A|B) = P(A)\ P(B|A) / P(B)}$$

Bayes' rule for events can be extended to Bayes' rule for random variables and their probability distribution functions. Bayesian inference estimates the posterior probability distribution function P(θ|data) of a set of parameters θ, given a set of observed data

$$\mathbf{P(\theta|data) = P(\theta)\ L(data|\ \theta) / P(data)}$$

with **P(θ)**, the prior probability that the parameters take the value **θ**, **L(data| θ)**, the likelihood function, i.e. the probability of observing the data given the parameters **θ** and **P(data)**, the probability of observing the data.

3.2.2 IPR+ predictive uncertainty calculation

Since the IPR+ index had already been presented and discussed with French river managers, the idea was to keep the method previously developed and simply add an uncertainty evaluation around the score. Consequently, the metric predictive models previously selected were implemented within a Bayesian framework (McCarthy 2007) using WinBUGS software (Spiegelhalter et al. 2003). In contrast to the former frequentist method, parameter estimations are no longer unique values but probability density functions (PDFs). For each metric, the PDFs of the model parameters were estimated using reference condition sites (Fig. 12). For a new test site, these models are used to predict metric PDFs in reference condition. Metrics are then transformed and aggregated to compute IPR+, as presented above. Consequently, the final IPR+ index is no longer a single value but takes the form of a probability distribution (Fig. 12). For convenience, scores of the metrics and IPR+ are described by the mean of their PDFs and their predictive uncertainty by the standard deviation of their PDFs (SD_U).

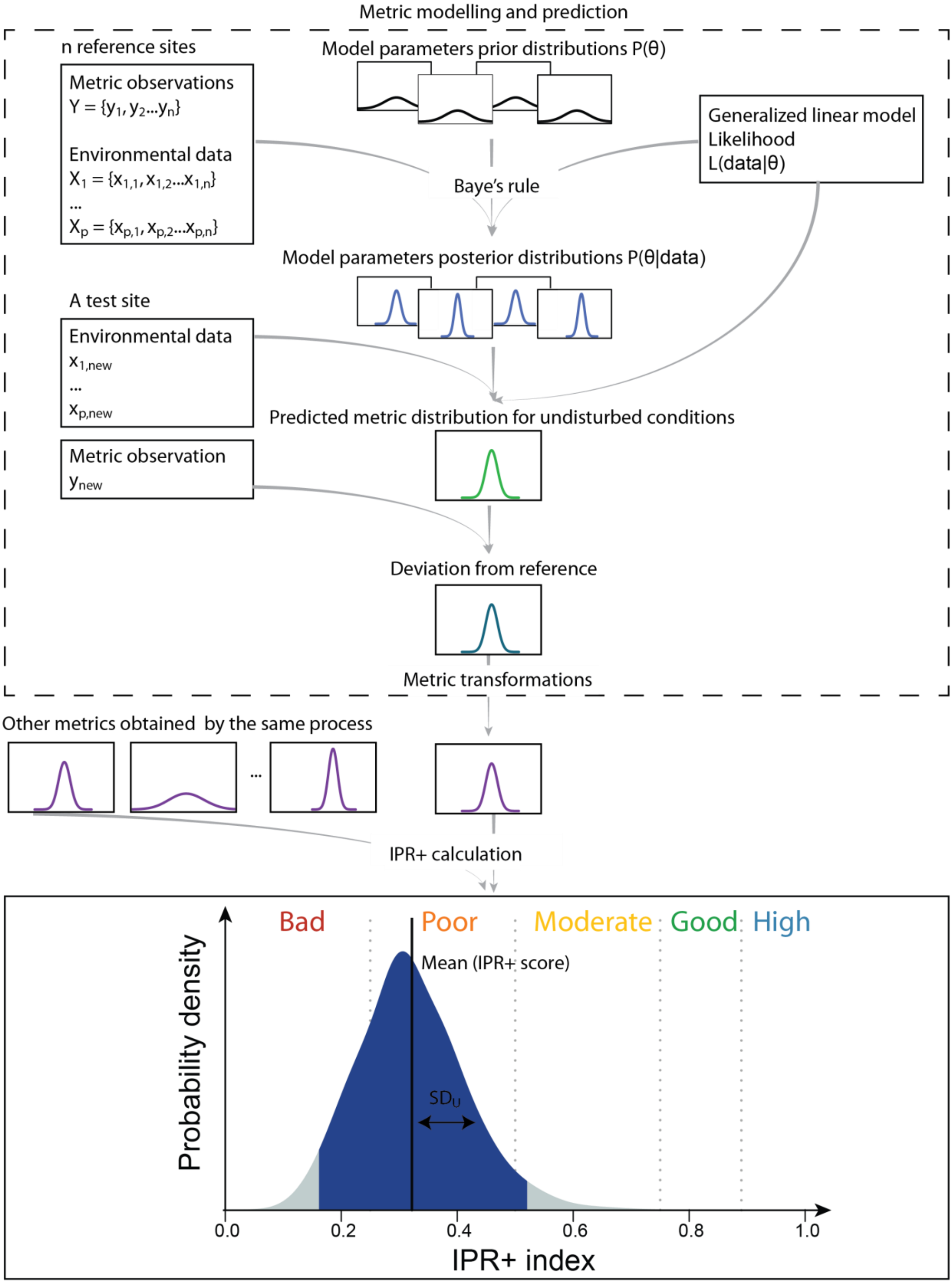


Figure 12. Calculation of the IPR+ predictive multi-metric index and propagation of the predictive uncertainty.

3.3 IPR+ SCORE INDEPENDENCE FROM THE ENVIRONMENT AND RESPONSES TO HUMAN PRESSURES

The effects of the natural and human sources of spatial heterogeneity on the IPR+ index score (mean of the PDF) and the way metrics are selected to enter the IPR+ calculation were examined using the IPR+ database (Fig. 1; ***P3*** p. 149).

Metrics involved in IPR+ calculation

Intolerant metrics enter more frequently into IPR+ calculation than other types of metrics. **Metric selection** does not seem to depend on the type of disturbances but on the **level of disturbance** and the **upstream or downstream position** (***P3*** p. 150). In quasi-undisturbed conditions, the metrics selected are those naturally highly represented in fish assemblages. Intolerance metrics were preferentially selected for upstream sites while S-LIMNO and S%-LIPAR were selected downstream in accordance with the longitudinal variation of fish assemblage structure (Logez et al. 2012).

Independence from the natural variability of environment

Analysis of variance procedures were used to test for the independence of IPR+ from physiographic variables (catchment area, 10-year mean stream power, 10-year mean air temperature, 10-year mean air temperature amplitude, catchment geological type and hydrological regime). **IPR + and underlying metrics did not vary significantly with physiographic variables** ($P > 0.05$), indicating that the part of fish community variability related to natural spatial heterogeneity of the habitats has been successfully removed.

Responses to anthropogenic disturbances

The responses of IPR+ scores to overall human disturbance, individual water quality, hydrological and morphological human-induced disturbances were analyzed using the same criteria as in section 2.3 (i.e. significance, intensity, sensitivity). In agreement with previous results (***P1***), the multi-metric **IPR+ index detects all the types of disturbance but is more sensitive to the combination of different river degradations** (Fig. 13a, b).

Interestingly, whereas previous bio-indicators usually demonstrated a lack of sensitivity to severe hydrological degradations alone (***P1***), IPR+ responds relatively well to the different types of alteration (Fig. 13 c, d, e). Nevertheless, IPR+ presents different sensitivities to human-specific disturbances. It responds to low morphological and water quality degradation but only to medium hydrological degradation.

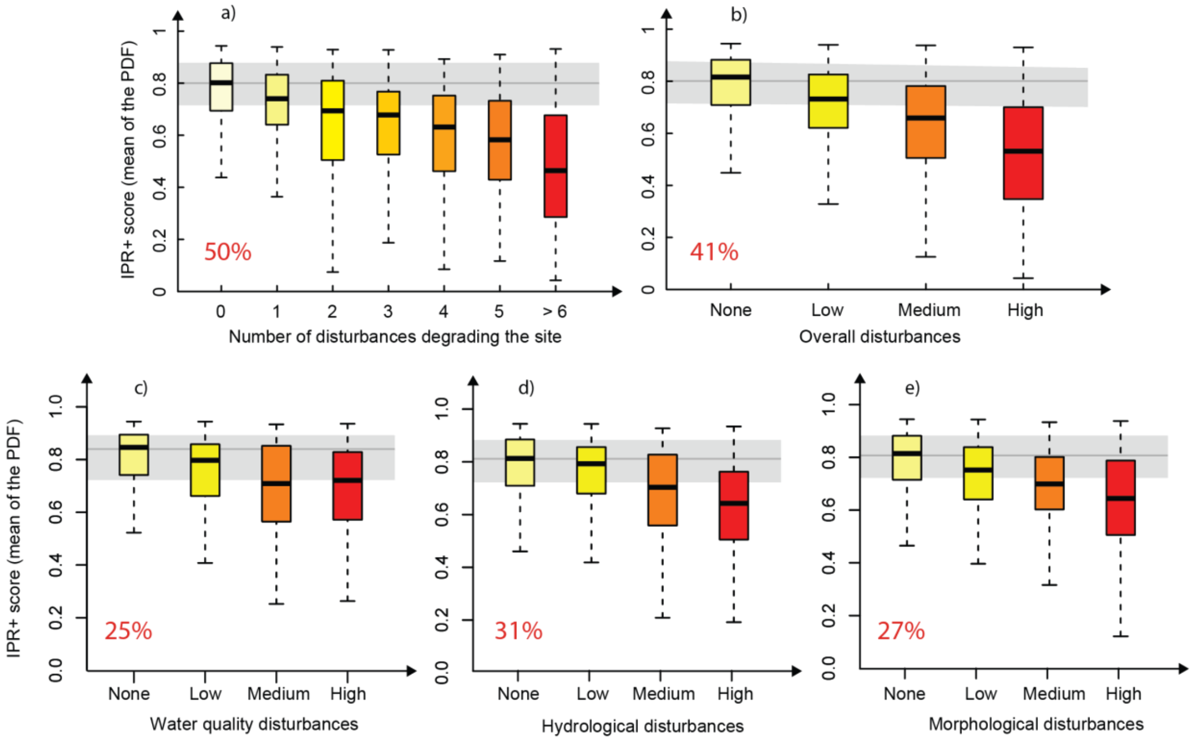


Figure 13. Responses of the IPR+ score to (a) the number of different types of human disturbances, (b) global, (c) water quality, (d) hydrological and (e) morphological human-induced degradations. All the responses were significant. Grey zone, the reference sites. Red percentages, discriminatory efficiencies quantify the intensity of the response (see section 2.3; Ofenböck et al. 2004).

Implications for bio-assessment

The method used to discard the effect of natural variability seems satisfactory since IPR+ scores were independent of physiographical features. This result validates the inclusion of moderately impacted lowland sites in the model calibration data set to account for natural variability in lowland-site communities.

Although IPR+ seems to be more sensitive to the accumulation of different degradations than to intensification of individual pressures, it appears suitable to detect the impacts of specific degradations including hydrological perturbations.

In addition, these results illustrate the value of selecting the most degraded metrics. The multi-metric index selects the metrics that are the most sensitive to the actual condition to reflect the ecosystem status, i.e. depending on the longitudinal position and level of degradations.

3.4 Two potential sources of uncertainty: predictive uncertainty and temporal variability (***P3*** and ***P4***)

Two aspects of the bio-assessment uncertainty were considered: (1) One is directly related to the prediction of metric values in the reference condition, i.e. predictive uncertainty, and (2) the other one is related to the inter-annual variations of the IPR+ score in absence of human-induced changes, i.e. natural temporal variability.

The first one was measured as the standard deviation of the IPR+ and metric score probability distribution (**SD_U**: probability distribution represented by the blue curves in Fig. 12 and 14) using the method described in section 3.2.2 and the IPR+ construction data set (*N* = 1654; Fig. 1, ***P3***). The second was measured as the standard deviation of IPR+ and metric scores over a 10-year period (1998–2007) for 183 sites (**SD_T**) (Fig. 1, ***P4***).

The amounts of the two types of uncertainty are compared. Then the influences of natural environmental features and human disturbances on predictive uncertainty and inter-annual variability are analyzed. Finally, the use of predictive uncertainty to define confidence in site classification is examined.

3.4.1 Predictive uncertainty vs. temporal variability

For the 183 time series, average predictive uncertainty of the IPR+ score was much larger (SD_U, 0.14) than inter-annual variability (SD_T, 0.07) (***P4*** p. 179). Consequently, in stable human disturbance conditions, over a 10-year period (1998–2007), only a few sites (e.g. Fig. 14c) show a significant inter-annual difference in IPR+ scores (no overlap of the PDF 2.5 and 97.5% percentiles between the two years in the period; ***P4*** p. 179). This result revealed that in the absence of changes in human disturbances, IPR+ index scores do not vary substantially between years (Fig. 14a, b) or IPR+ index scores do vary but individual scores are too uncertain to distinguish scores between the two years (Fig. 14d).

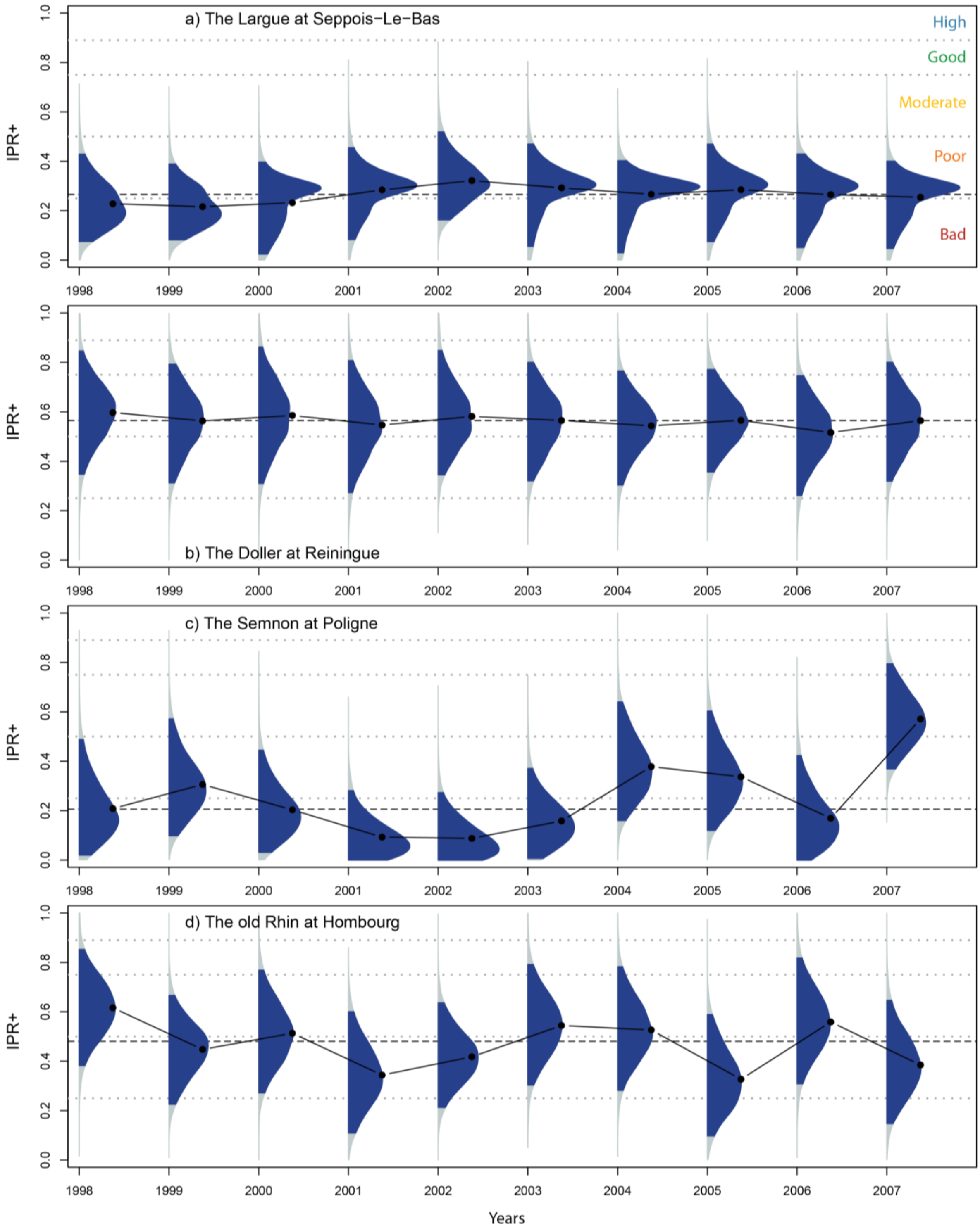


Figure 14. Inter-annual and predictive variability of four individual sites. PDFs of the IPR+ index for 10-year time series. Blue area, 95% probability of the IPR+ score. Grey area, the other 5%. Black dashed horizontal line, 10-year median of the IPR+ scores. Black dots linked with black line, IPR+ scores, i.e. mean of the PDFs. Grey dotted lines, the five ecological status classes.

3.4.2 Influences of the longitudinal gradient and human pressures

Predictive uncertainty

IPR+ **predictive uncertainty is similar upstream and downstream** (***P3***), suggesting that models have the same predictive power for upstream and downstream sites. Metric SD_U are generally higher than IPR+ SD_U (***P4*** p. 179).

The bell-shaped relationship observed between IPR+ SD_U and IPR+ scores (Fig. 15) shows that **index uncertainty is smaller and less variable for highly impacted and non-impacted sites** (bad and high status) **than for middle-range sites** (moderate status). This pattern was confirmed looking at the effect that human-induced disturbances had on SD_U. IPR+ scores were significantly less uncertain for undisturbed and strongly disturbed sites than for intermediate levels of disturbance (Fig. 16a, b; ***P3***).

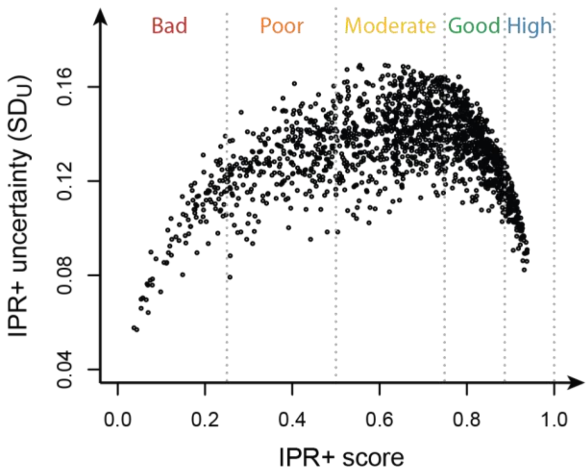


Figure 15. IPR+ predictive uncertainty versus IPR+ score

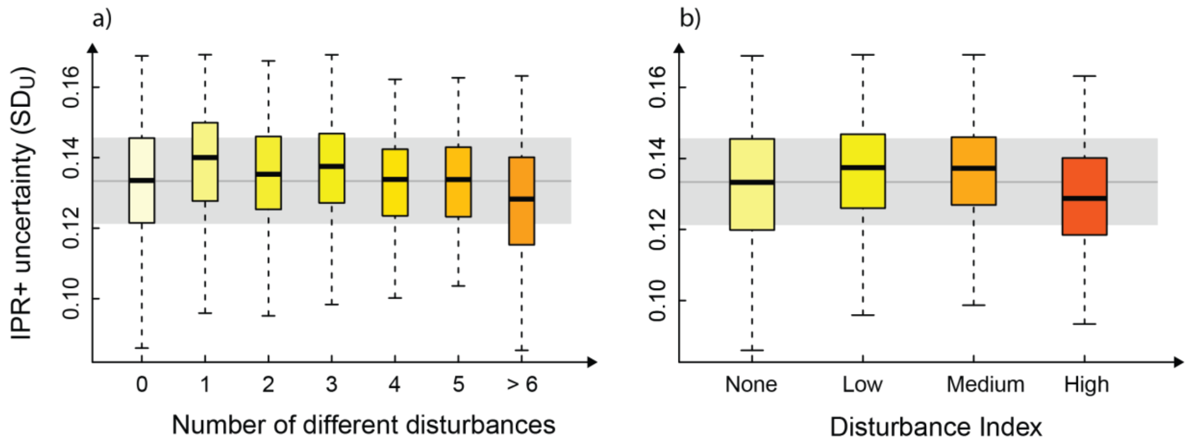


Figure 16. IPR+ predictive uncertainty and levels of human disturbance

In addition, this pattern might be amplified by the transformations necessary to acquire an EQR. Scores are constrained between 0 and 1, with 0 the most degraded sites used to develop the IPR+ index and 1 the best reference sites of the dataset However, in reality the test site could have an extremely low (some probability of being < 0) or very high ecological health (better than reference sites, some probability of being > 1). The piece-wise transformation constrained these site scores between 0 and 1 and mechanically their probability of being 0 or 1, respectively is very high.

Temporal variability

Year-wise Pearson correlations of the IPR+ scores were calculated for the 183 sites over the 1998–2007 period. The IPR+ scores gave **consistent evaluations over the 1998–2007 period** (r > 0.74; ***P4*** p. 179). In agreement with Collier's work (2008) on the Macroinvertebrate Community Index, while some of the metrics were more stable through time in upstream than in downstream sites, IPR+ temporal variability was similar upstream and downstream (***P4*** p. 180).

In addition, an increase in human disturbance intensity weakens the temporal stability of the fish assemblage structure (section 2.2; Collier 2008; ***P4***) and consequently affects the temporal stability of bio-assessment indicators relying on functional metrics (Fig. 17). This result is in agreement with several studies showing that indicators were less stable over time in degraded situations than in slightly disturbed conditions (Fore et al. 1994; Ross et al. 1985).

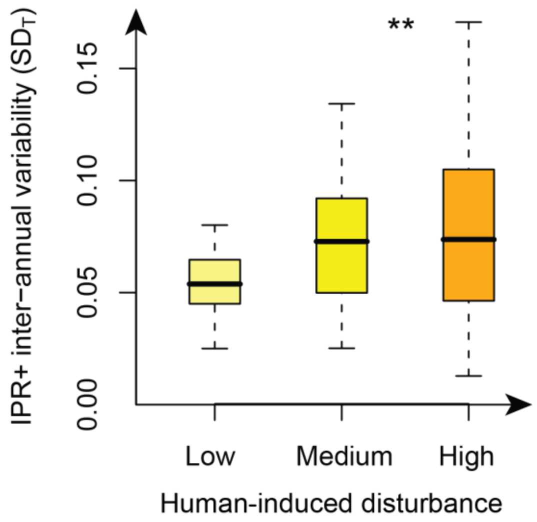


Figure 17. IPR+ inter-annual variability and longitudinal gradient (p-value < 0.01)

In addition, probably due to density dependence phenomena (Jenkins et al. 1999; Kaspersson & Höjesjö 2009), the compositional metric N%-Trout (relative abundance of trout juveniles) was globally less stable over time than the other functional metrics (***P4*** p. 179). This result is in accordance with previous studies suggesting that functional assemblage structure is more stable between years than taxonomic composition and species abundance (Bêche et al. 2006; Fransen et al. 2011).

Finally, these results suggest that the index's lack of inter-annual stability could be an additional evidence of the system's severe degradation.

3.4.3 Confidence in site classification

For decision-making purposes, the chance that the IPR+ score of a given site is contained within status classes can be derived from the SD_U, i.e. confidence in site classification. The analysis of confidence in site classification attempts to answer the following question: "Are the sites truly above or below the class boundary?" Consistent with Clarke et al. (1996), the confidence in class status increases with the score distance to the class limit (Fig. 18). Astonishingly, it seems that there is a shift in the classification confidence for the "good" status class. This is probably the result of the piece-wise transformation constraining the score to acquire EQR varying from 0 to 1 with the average score for the reference sites at 0.8 (Pont et al. 2012).

The mechanical cause of this result explains why confidence in the classification could be low even if the uncertainty around the score is relatively low. Knowing this, different rules could be applied to decide whether or not a site truly has good status. For instance, Ellis (2007) proposed the benefit-of-the-doubt rule (a site has "good" status if it has less than 95% confidence of being in the "bad-poor-moderate" classes).

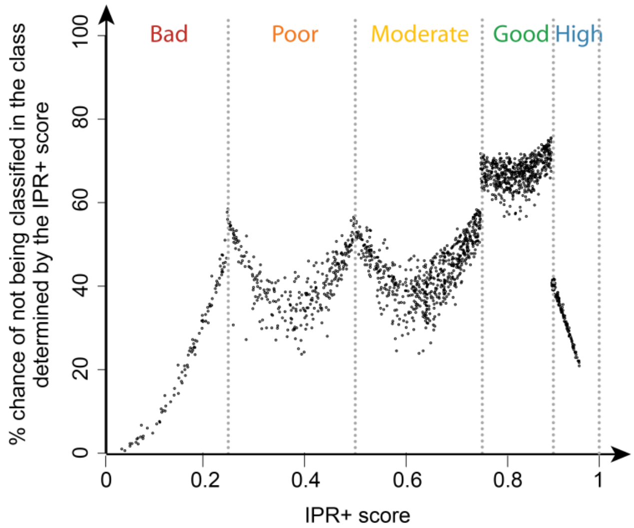


Figure 18. Confidence in site classification versus IPR+ score

3.4.4 Implications for bio-assessment

In this section, two sources of uncertainty in decision-making were examined. The superiority of SD_U over SD_T does not mean that inter-annual variation of IPR+ does not exist but that considering the predictive uncertainty, the accuracy of the predictions could be too low to significantly detect them.

However, the relatively low SD_T confirms the value of the multi-metric index based on functional metrics given that the latter are more stable from one year to another than compositional metrics. In addition, metric and index stability might be improved by taking short-term climatic variables into account in the models.

This result implies that excessive predictive uncertainty might make it difficult to detect finer spatial and temporal variations in fish community integrity. It emphasizes the need to reduce uncertainty to improve the power of biological indicators to detect human disturbance impacts. This first step quantifies this uncertainty and the following step is designed to find ways to minimize it. Inter-annual variability might be reduced by the use of sampling repetitions instead of a 1-year single sample to evaluate the river sites. The set of informative prior PDFs and other more sophisticated Bayesian methods could be some leads to pursue to diminish predictive uncertainty (Gelman et al. 2004).

These two sources of uncertainty are generally greater for individual metrics than for the IPR+ index. These results indicate that for a multi-metric index the choice and combination of particular metrics may reduce the index's temporal variability and predictive uncertainty. The aggregation of metrics into multi-metric indices may lessen the predictive uncertainty and the temporal variability of metrics related to environmental stochasticity (Bêche et al. 2006) and accordingly enables a more reliable assessment of stream ecological conditions over several years.

Finally, these results strengthen the idea that biological indicators should be associated with uncertainty measures. In addition, even for very little uncertainty, the mechanical effect of classification could decrease the confidence in site status class as the site score becomes closer to a class limit (Clarke et al. 1996).

3.5 CONCLUSIONS

(1) As bio-indicators are devoted to evaluating the ecological quality of water bodies, their **uncertainty and uncontrolled variability have crucial implications for water management decisions** and need to be acknowledged.

(2) The **Bayesian method developed** herein seems **appropriate to build a multi-metric index associated** with a measure of **predictive uncertainty**. The resulting IPR+ scores are independent from natural variability and are sensitive to human-induced disturbances. In particular, they are affected by the accumulation of different degradations and specific degradations including hydrological perturbations.

(3) In absence of changes in human disturbance, **evaluation using IPR+ scores** is globally **consistent through time** over a 10-year period.

(4) In absence of changes in human disturbance, **IPR+ predictive uncertainty generally overcomes IPR+ score temporal variability**.

(5) IPR+ predictive uncertainty and temporal variability are **similar upstream and downstream**.

(6) While **IPR+ uncertainty is higher for mid-range** sites (moderate status), **inter-annual variability is higher for sites severely degraded** by human pressures.

(7) Predictive uncertainty and **temporal variability** are globally **lower for the IPR+ multi-metric index** than for **underlying metrics**, and especially for the compositional metric (trout juveniles).

To our knowledge this is the first time that a multi-metric index has integrated the uncertainty associated with establishing reference conditions for present and future climatic conditions. In light of these results, Bayesian modelling seems a suitable method to acquire an explicit measurement of the uncertainty around reference conditions essential for decision-making and more generally to answer bio-assessment issues. This new methodology is relatively generic and could be extended to other biological groups and over larger spatial extents.

These studies confirm the advantages of the multi-metric index based on ecological and biological functional metrics for the assessment of river ecological health.

Finally, two sources of uncertainty in river management decision-making were examined as a first step in IPR+ uncertainty recognition: the temporal variability and the predictive uncertainty of the IPR+ score. The error related to the sampling method and effort variability should also be recognized to provide a complete view of the uncertainty around the index score. In future steps, it would be advantageous to find ways to minimize uncertainty. Sampling repetitions instead of single-sample evaluation as well as advanced Bayesian methods could be one of the leads for this challenging quest.

4. FROM BIO-INDICATION TO RIVER RESTORATION (***P5***)

Since 1990, increasing public awareness of river degradations and, more recently, society's crucial need for river ecosystem services (Baron et al. 2002) have led to a surge in river restoration projects throughout Europe and elsewhere, mainly North America (Fig. 19) (Bernardt et al. 2005; Palmer et al. 2007).

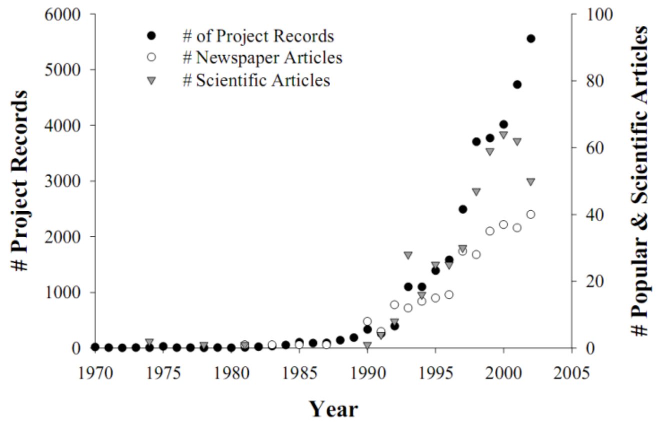


Figure 19. Number of US restoration projects, newspaper and scientific articles (from Bernardt et al. 2005).

Biological indicators were originally developed to detect the impacts of human degradations (***P1***, ***P2***, ***P3***) in order to prioritize the measures of river quality improvement (CIS 2012). Once restoration measures have been implemented on targeted sites, the degree of success of the measures undertaken has to be determined, i.e. the effects on river biotic and abiotic features. However, restoration measure outcomes have been poorly documented, collecting and making accessible this critical information is crucial to improving restoration designs (Palmer et al. 2007).

Weir and small dam removal is one of the common measures implemented to improve river ecosystem health. Indeed, dams and weirs have major impacts on river biota (***P1***; ***P2***; Baxter 1977; Petts1984; Stanford et al. 1996). By removing small dams, the longitudinal connectivity and the natural free flowing functioning of streams and rivers are restored and positive impacts on the river biota are expected (e.g. Shuman 1995; Kanehl et al. 1997; Born et al. 1998; Leaniz 2008). However, due to the dramatic changes induced, some authors have suggested that weir removal may be considered as ecological disturbances in their first stage (Stanley & Doyle 2003), mostly due to the release of large amounts of sediment downstream. Consequently, it is not certain that the simple reversal of degradation provides the desired and anticipated ecological effect.

In 2010, as part of the European WISER project, among three measures studied in the project, I examined the impact of weir and small dam removal (***P5***). The aim was to ascertain what has and what has not been learned over the last few decades on the effects of this measure on river ecosystems (Feld et al. 2011).

In this section, **(1)** a **conceptual model** establishing the relationships between weir removal and abiotic and biotic river characteristics is presented and **(2) the lessons** learned **from previous case studies** are reviewed.

4.1 WHAT HAS BEEN ACHIEVED BY REMOVING SMALL DAMS AND WEIRS?

The DPSIRR conceptual model

In the WISER project, a **recovery** component was added to the conceptual DPSIR (**Driving force, Pressure, State, Impact, Response**; Fig. 2) framework to describe the effects of societal responses (i.e. restoration measures) on river ecosystems (abiotic and biotic elements). The resulting **DPSIRR** conceptual models structure and illustrate current knowledge on observed relationships of restoration measures (cause) and river abiotic and biotic conditions (effect). This conceptual framework links socio-economics with ecology and identifies the **well-known** cause–effect chains, **knowledge gaps** and **lack of scientific evidence**.

In 2010, to address this issue, evidence of weir removal effects on abiotic and biotic characteristics (algae, macrophytes, macroinvertebrates and fishes) were compiled from restoration studies reported in peer-reviewed and grey literature and structured within a conceptual framework. The resulting conceptual model (Fig. 20) was based on a review of 31 restoration studies and five general ecology papers (***P5*** p. 199). Most of these studies sampled sites of one to a few kilometres long and compared conditions before and after weir removal. However, only half of the studies applied a full BACI design (Before-After-Control-Impact; Friberg et al. 2011). The main assumption of the weir removal response/recovery conceptual model is that hydrological and ecological connectivity are major determinants of physical habitat in streams (e.g. flow and habitat diversity, sediment transportation and size, water quality), which in turn are major determinants of river biota characteristics (e.g. composition, diversity, biomass, function) (Bunn & Arthington 2002).

Weir removal effects on abiotic factors

Weir removals re-establish the hydro-morphological and ecological connectivity of streams and rivers (e.g. Poff et al. 1997; Gregory et al. 2002). This transformation results in a shift from lentic to lotic river systems upstream, leading to the reservoir sediment release and a pulse of fine sediment downstream (e.g. Randle 2003; Pollard & Reed 2004). Upstream, fine materials (e.g. sand, silt and mud) erode and uncover coarser substrata (e.g. gravel, pebble and cobbles), which enhances the overall habitat diversity (Kanehl et al. 1997; Born et al. 1996) and water quality (e.g. dissolved oxygen; Hill et al. 1994). In addition, other studies commonly reported an increase in flow diversity (e.g. Bushaw-Newton et al. 2002) and decreased water temperature upstream from the former dam (e.g. Kanehl et al. 1997; Hill et al. 1994). However, downstream from the former dam, the flush of former impoundment fine sediments increases turbidity, at least temporarily (e.g. Chaplin 2003). In addition, fine sediment movement could have strong abrasive effects (e.g. Chaplin 2003; Cheng & Granata 2007) and contaminated sediments could be brought downstream (Bednarek 2001). Very often, not all sediments accumulated upstream from the dam are flushed downstream and sediment legacies remain in place for long periods of time.

Consequences on the biota

It is important to note that only the proven effects, thoroughly described in the literature, were integrated into the model. Studies mostly examined the responses of macroinvertebrates and to a lesser extent macrophytes and fishes to dam removal, while benthic diatoms were rarely considered. Changes in water quality such as sediment size and turbidity affect benthic algae and sensitive macroinvertebrates (mainly Ephemeroptera-Plecoptera-Trichoptera taxa) (e.g. Baattrup-Pedersen & Rijs 1999; Bushaw-Newton et al. 2002; Orr et al. 2006). Macrophyte communities seem mainly influenced by changes in channel morphometry (depth, width) and connectivity (e.g. Shafroth et al. 2002). Fish communities appear to be mainly sensitive to the re-establishment of longitudinal ecological and hydrological connectivity that restore and make accessible habitats suitable for some fish species' life stage (e.g. gravel bars upstream for salmonid species reproduction) (Iversen et al. 1993; Poff et al. 1997; Leaniz 2008). However, the flush of fine sediment downstream could also have the opposite impact on the biota and consequently particular care should be taken concerning this aspect (e.g. for fish species weirs should not be removed close to the spawning season).

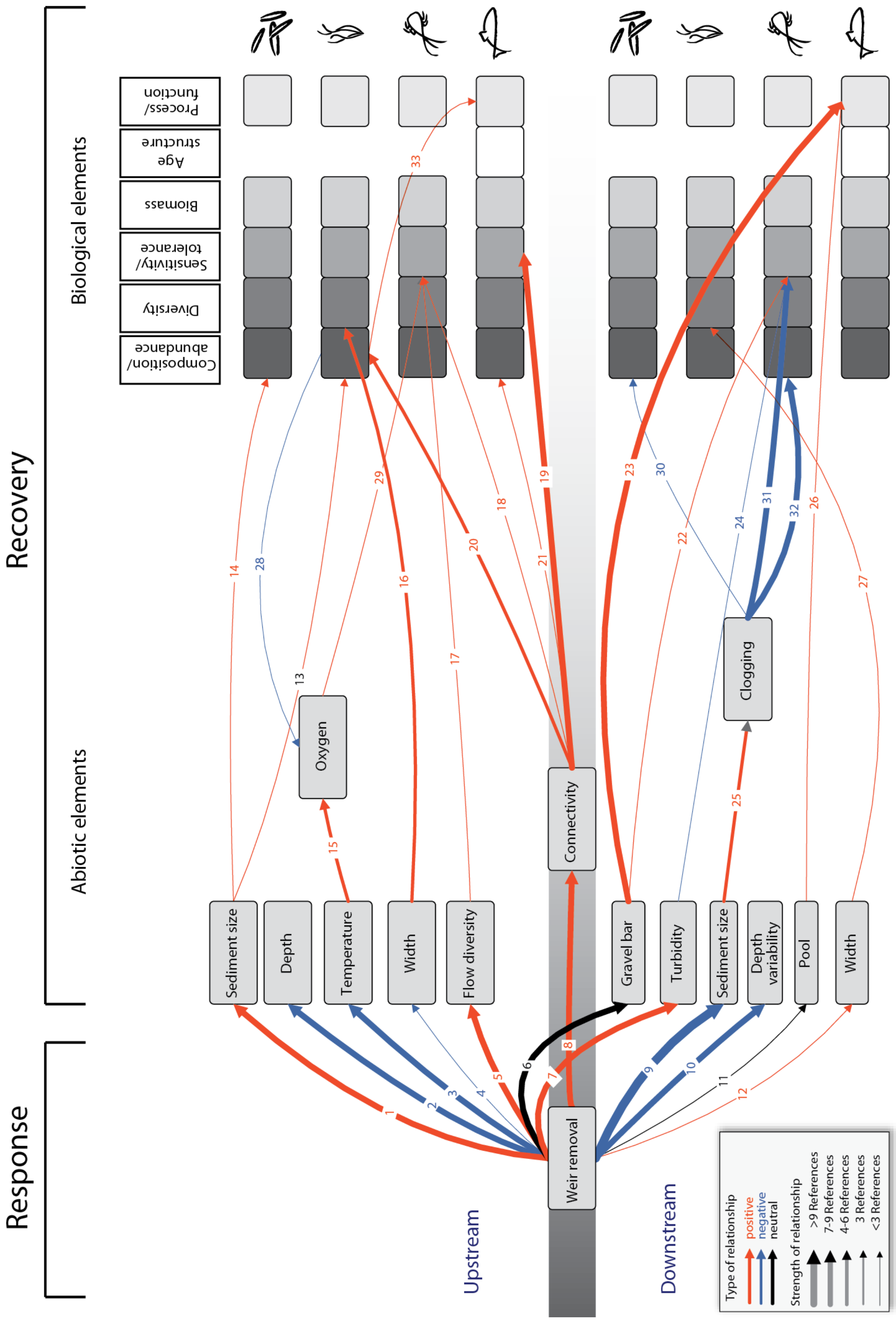


Figure 20. Conceptual models of response/recovery chains in weir removal. Each link refers to a relationship referenced in the restoration literature reviewed (see ***P5***, Appendix B, p. 199). Relationships may be either positive (red arrows), negative (blue arrows) or equivocal/ambiguous (black arrows)

Time scale of the recovery

After weir removal, each river characteristic evolves following a specific time scale (Doyle et al. 2005). While some of them take years to centuries to recover, others recover in days to months. Return of free-flowing conditions as well as the effects on water temperature and longitudinal connectivity is immediate after the removal of weirs. In contrast, biological recovery is expected to occur once fine sediments have been transported farther downstream and geomorphological processes are restored, which could require several years to decades after weir removal (e.g. Thomson et al. 2005). These effects largely depend on the quantity of sediments accumulated above the barrier, water velocity, the gradient of the riverbed, and also on the specific technique used to remove the dam (Bednarek 2001).

Although full recovery may take decades (Bednarek 2001), for the studies reviewed the monitoring period ranges from 3 months to several years. Consequently, the time-scale and trajectory of recovery after weir removal remain partially unknown and long-term monitoring data are still required.

4.2 KNOWLEDGE GAPS AND LEADS FOR FUTURE RESEARCH

This review of the literature shows that during the last decade substantial efforts have been made to investigate the processes initiated and restored by weir removal. Nevertheless, the results were mostly **qualitative** and there is still a **lack of quantitative measurements** of the effects that are necessary to predict processes in order to estimate and predict impacts of weir removal on river communities.

In addition, the literature provided little information on the **effectiveness** of weir removal restorations. This was rarely measured and the criteria were usually vague. In most cases, negative impacts of weir removal were rather short term (e.g. increase in suspended sediments), while the assumed beneficial changes were likely to act over the longer term (e.g. increase in flow diversity, connectivity); the natural free-flowing state of the river was always regained, whereas recovery of the biota following this habitat shift was more uncertain. Furthermore, the effects of restoration could be highly variable depending on the hydrologic nature of the river (Chaplin 2003). As a result, the effectiveness of a dam removal is likely to vary widely among systems and depends both on temporal and spatial scales of the restoration scheme.

4.3 CONCLUSIONS

(1) Whereas the natural free-flowing state is immediately regained after weir removal, hydro-morphological, water quality and biota components could take **months to decades to recover**.

(2) Relatively short-term negative impacts of the weir removal are mainly caused by an increase in suspended sediments downstream from the former impoundment.

(3) However, **long-term beneficial impacts** are observed, such as increases in flow diversity, connectivity or upstream water quality.

(4) Recovery of the **biological communities** is **more uncertain** and might take **longer** than the other components.

(5) Monitoring weir removal rarely exceeded a few years and the full **recovery time-scale is still partially unknown.**

(6) The **effectiveness** and success of recovery was rarely measured and **quantitative measurements of weir removal effects** are **lacking.**

These results highlight the need for quantifying the long-term effects of restoration on river ecosystems and the importance of monitoring river characteristics before and after weir removal over a long period. In the previous sections, biological indicators responded to human-induced degradations (***P1***, ***P2***, ***P3***). To measure and possibly predict the degree of success of the measures undertaken, the extent to which biological indicators detect improvements of the river conditions (intensity and sensitivity of the responses) needs to be assessed.

Furthermore, such conceptual models can provide useful tools for planning more effective river improvement measures and for identifying leads for future research in restoration ecology. They should be updated as restoration knowledge is improved.

5. General discussion and Perspectives

This thesis addresses several aspects and assumptions of river biological assessment as well as certain limits of biological indicators and suggests solutions and recommendations for future developments and uses of biological indicators.

Three main objectives shaped this work:

(1) Testing some of the assumptions that influence the choices of biological groups and community descriptors to assess the integrity of river ecosystems.

(2) Using an existing index (IPR+) to develop a method to account for biological indicators' uncertainty related to the prediction of reference conditions and addressing two aspects of uncertainty in bio-assessment: the predictive uncertainty and the temporal variability of bio-indicators.

(3) Reviewing what has been learned about the effects of improvement measures on river communities using a conceptual framework through the example of weir removal, in an attempt to advance towards restoration.

The different analyses highlight:

(1) The suitability of multi-metric indexes and functional metrics (***P1***, ***P3***, ***P4***) for river bio-assessment.

(2) The predominant role of physiographic factors (compared to human pressures) in shaping biological community compositions (***P2***).

(3) The variation in bio-assessment uncertainty (predictive uncertainty of the IPR+ index and temporal variability) with human pressure levels and natural environmental conditions (***P3***, ***P4***).

(4) The different responses of biological assemblages and related biological indicators to human disturbances (***P1***, ***P2***).

(5) Hierarchical impacts of human pressures on biological assemblage composition and biological indicators (***P1***, ***P2***) and the existence of complex effects shaping river communities, i.e. interaction effects between different types of human pressures (***P1***) and human pressures and environmental factors (***P2***).

(6) The equivocal effects of weir removal on river communities (***P5***).

What are the relevant indicators of the impacts of human disturbances?

The results of ***P1***, ***P3*** and ***P4*** show that the responses of individual indicators based on the same river biological group could be highly variable. The selection of bio-indicators should not only focus on the BQEs, but also on the nature of the metric (i.e. underlying processes and units). Three types of bio-indicator were compared: the functional metric, the compositional metric and the multi-metric index. Indexes and functional metrics show the strongest responses to human disturbances. Indexes respond significantly to the different types of human pressures. Functional metrics generally detect lower levels of human disturbances than compositional metrics. In addition, the only compositional metric (relative abundance of trout juveniles) integrated into the IPR+ index is globally less stable over time than the other functional metrics (***P4***). This result is in accordance with previous studies suggesting that functional assemblage structures show less spatial and temporal variability than taxonomic composition and species abundance (Bêche et al. 2006; Fransen et al. 2011). Ecological and biological functional traits are better adapted to large-scale approaches (Statzner et al. 2001; Lamouroux et al. 2002) and are able to integrate more general phenomena than compositional metrics (Dolédec et al. 2006). Furthermore, predictive uncertainty and temporal variability are globally lower for the IPR+ index than for underlying metrics.

These studies confirm the **advantages of multi-metric indexes based on ecological and biological functional metrics for the assessment of river ecological health**. In addition, ***P3*** illustrates the value of the differential selection of the most degraded metrics at each site to adapt the index to different conditions (longitudinal position and level of degradation). At each site, the most sensitive metrics to the actual condition are used to reflect the ecosystem status.

Perspectives

Further research is still needed to determine relevant functional traits describing the integrity of river communities, particularly for macrophytes. Moreover, food web ecology provides means to quantity the impacts of human pressures on river ecosystem structure, function and stability by considering the role of complex trophic interactions structuring river communities (e.g. Friberg et al. 2011; Thompson et al. 2012). Defining and including metrics and indicators based on such integrative measures might enhance the ability to detect human impacts on river ecosystem functioning.

Taking into account bio-assessment uncertainty

The variance partitioning analyses (***P2***) and the standardization transformations (***P1***) confirmed the major role played by natural spatial variability in abiotic features (e.g. river size, climate, geology and hydrological regime) in structuring river biological communities. In addition, the analysis of the inter-annual variability of fish assemblage composition in absence of change in human disturbance illustrated the effects of natural disturbances (e.g. drought or floods) and stochastic biological processes on river communities. Consequently, this variability (noise) should be distinguished from the effects of human-induced changes (signal) so as not to bias the assessment of the ecological integrity.

The use of statistical models to predict community characteristics in quasi-absence of human disturbances, i.e. the predictive reference condition approach (Oberdorff et al. 2001; Pont et al. 2006; 2007), appears to be appropriate to eliminate the variability in metrics and biological indexes related to natural spatial variability (***P2***, ***P3***). The natural temporal variability of indexes and metrics related to climatic hazards could be limited by including climatic variables (temperature and precipitation) in the models (Linke et al. 1999; Mazor et al. 2009).

However, statistical models in essence simplify reality and samples of river communities are more or less representative of the actual community depending on the sampling method and effort. For these two reasons, the recognition of spatial and temporal natural variability could be incomplete and could bias the prediction of metrics in absence of human disturbances and thus the assessment given by the multi-metric index, i.e. predictive uncertainty. In addition, the models did not recognize the fact that responses of aquatic communities to abiotic features are likely to change over time through stochastic biological processes (e.g. extinction, recruitment and biotic interactions). Although the effect of the sampling effort variation was minimized by considering only samples with reasonable sampling efforts, in combination with community and population dynamics, this variability might explains the temporal variability in the indexes and metrics observed in absence of change in human disturbances. These two sources of uncertainty related to the incomplete recognition of natural spatial and temporal variation in river communities are a potential risk for decision-makers and need to be acknowledged.

Temporal variability – In the absence of changes in human disturbance, evaluations using IPR+ scores were globally consistent with time over the 10-year period (***P4***). The inter-annual variability of the IPR+ index score is higher for sites that are severely degraded by human pressures than for less disturbed sites. This is in agreement with the results of previous analysis (section 2.2) and studies (Fore et al. 1994; Ross et al. 1985; Collier 2008) showing that the inter-annual variability of fish community structure increases with the human disturbance level. Schaeffer et al. (2012) suggested that redundancy of species traits might mitigate the impacts of human pressures on river ecosystem functioning. Species loss related to human disturbances limits these compensatory forces and could destabilize the functional structure of communities (Franssen et al. 2011). Assuming that **human disturbances destabilize river community structure**, the measurement of the inter-annual variability of the index score could be used as additional evidence of the degradation of the ecological integrity of the river.

Perspectives

As global climate change progresses, the natural temporal variability of bio-indicators might increase. In this context, it is particularly relevant to integrate mid-term climatic data such as air temperature and precipitation into the models. This partly accounts for this inter-annual variability and predicts how reference conditions may evolve for future climatic conditions. Monitoring the evolution of reference conditions following climate change is therefore essential to be able to measure and recognize its effects on river biota (Nichols et al. 2010). Once the different scenarios of future reference conditions have been determined, future river management objectives should be adapted.

However, since climate change will modify river functioning, it is quite possible that the impacts of the same human pressures on the biota could be either increased or mitigated (i.e. interaction effects between human pressures and environment). In the longer-term (30–50 years), the metrics previously selected might no longer be representative of French river ecosystems (modifications of the regional species pool or of the natural environment) and adapted to detecting impacts of future human pressures (Logez & Pont 2012). Consequently, in the future, beyond a certain point of climate modifications, it would be necessary to revise bio-indicators and redefine reference conditions for assessing the ecological conditions of rivers.

Predictive uncertainty – The Bayesian method developed appears to be relevant to account for the predictive uncertainty of a multi-metric predictive index (***P3***). The resulting IPR+ scores are independent of natural variability and are sensitive to human-induced disturbances. In particular, the IPR+ is impacted by the combination of different degradations and specific degradations including hydrological perturbations.

Perspectives

This methodology could be used to build bio-indicators based on other biological groups and metrics while recognizing predictive uncertainty. Also, using corresponding data sets to calibrate the models, it could be extended to other spatial areas presenting comparable river functioning and human pressure categories.

The analysis of predictive uncertainty reveals that the index scores are comparable upstream and downstream but are more uncertain and more variable for sites with middle-range scores (moderate status class, ***P3***). This result might be partly explained by the commonly used piece-wise transformation to acquire EQR. It would be valuable to understand the factors structuring the high variability in predictive uncertainty for moderate-status sites. A complementary analysis attempting to relate this variation in uncertainty with physiographic variables (e.g. distance to the source, geology) has been conducted, but the results did not highlight a relevant pattern.

Unimpaired lowland sites were rare in the national French data set, but it was necessary to find a way to assess them. There are different approaches to defining reference conditions (Stoddard et al. 2006). Some authors advocated that reference conditions should be a balance between truly pristine sites and acquiring enough sites to allow meaningful statistical relationships (Pardo et al. 2012). Developing the IPR+ index, minimally impacted to moderately impacted sites were integrated to predict the reference conditions of lowland sites (***P3***). In addition, comparable predictive uncertainties were observed upstream and downstream. This aspect emphasizes the subjectivity of the definition of reference conditions and the need to define comparable criteria for the selection of reference conditions throughout European countries. In addition, since they are designed to assist river management, they should seek to meet achievable rather than idealistic objectives.

From uncertainty recognition to index accuracy improvement – In my point of view, the acknowledgement of the uncertainty around the index is clearly indispensable for decision making, but it is also a first step in improving bio-assessment reliability. In the case of the IPR+ index, the analyses of two possible causes of bio-assessment uncertainty reveal that predictive uncertainty is generally much higher than the uncertainty caused by inter-annual variability (***P4***). Furthermore, the test of inter-annual variability using predictive uncertainty implies that excessively high predictive uncertainty might make it difficult to detect finer spatial and temporal variations in fish community integrity. Therefore, it is a priority to attempt to minimize the predictive uncertainty in order to improve the reliability of IPR+ and its power to detect human pressures.

Perspectives

Several levers can be pulled to reduce predictive uncertainty. A first one is the quality of the data used to calibrate the models and to predict the metrics in reference conditions.

The sampling efforts and methods should be appropriate so as to acquire representative observations of the river communities. Using several samples from the same site might reduce the amount of uncertainty related to sampling errors. Several studies are currently investigating these aspects in French methods (Archaimbault Virginie, personal communication; Tomanova et al., in revision, the extent of sampling needed to acquire representative observation of large river fish communities), but further research and work are needed to determine how to increase data quality. The accuracy of the measurements of abiotic factors as well as the determination of human pressure levels could play an important role in controlling predictive uncertainty as well.

In addition, the taxonomic level at which individuals are identified should be consistent with the organisms' life traits and strategies, particularly macroinvertebrates. In the French national data set, macroinvertebrates are mostly identified at the genus or family level. Nevertheless, the lesser pertinence of the relationships between abiotic factors and taxa for macroinvertebrates than for fish in ***P2*** might be partly the result of a coarse level of identification. Genera and families could include species that are completely different in terms of biological traits and ecological preferences and consequently blur the responses of the communities to human and natural factors.

Previous knowledge on metric relationships with physiographic variables acquired with the data set covering comparable rivers could be used to set the informative model parameter prior to distributions and might reduce predictive uncertainty. More sophisticated methods such as hierarchical Bayesian models could also be a lead to explore for reducing predictive uncertainty (Gelman et al. 2004).

Moreover, the temporal variability of the index could be reduced by selecting metrics that are less variable over time (Hughes et al. 2004). For instance, the analysis of inter-annual variability of metrics showed that functional metrics are more stable than compositional metrics. Some functional traits seem more stable through time than others (e.g. for fish N%-RHPAR is more stable than S-TOL; ***P4***). More generally, using both minimal temporal variability and uncertainty as criteria for metric selection could greatly improve the quality of the assessment of the resulting multi-metric index.

Predictive uncertainty as a decision-making tool, confidence in classification

The use of probability density functions to evaluate the uncertainty around the index score is particularly well suited to river management decision-making as it is easily transformed into confidence in site classification (***P3***). Once the probability distribution and confidence in site classification have been determined, different rules can be used to decide whether or not a site is truly in good ecological health. In my opinion, this definition belongs to water managers, but it worth noting that the same probability distribution could lead to opposite results and decisions depending on the rule used. These rules should be defined in agreement with the final objectives of river management. If the final goal is to save money, one will choose the benefit-of-the-doubt rule, i.e. all the sites that have less than 95% of chance to be in worse than good condition are considered in good condition; in other words, all the sites that have 5% or greater chances of being in good condition are considered in good condition (Ellis 2007). The opposite rule will be to consider that only sites that have 95% or greater chances of being good are truly in good ecological "health". Intermediate alternatives will be to let the score (mean of the distribution) decide the ecological condition or simply which side of the Good/Moderate class limit includes the greatest share of the distribution.

Perspectives

In a further step, the evaluation of bio-indicators should test the implications of the different rules on the national evaluation.

Which BQEs should be used to reflect river ecological health?

The comparison of the responses of different river biological compartments to human disturbances (***P1***, ***P2***) shows that depending on their life traits and strategies, organisms and related bio-indicators respond differently in terms of intensity (i.e. is the response easily detectable?) and sensitivity (i.e. from which level of disturbance is the response detectable?). Benthic diatom and macroinvertebrate metrics appear to be the most sensitive to the degradation of the overall condition of the river. Benthic diatom and macrophyte metrics are highly sensitive to water quality degradation. Fish metrics generally have the highest intensities of response but lower sensitivity to human pressures (***P1***; Johnson & Hering 2009). These differences in sensitivity may be partly related to fishes longer life cycles and migratory capacities to avoid and survive disturbances in contrast with more sedentary short-life-cycle organisms such as benthic diatoms. In addition, fish metrics are the most impacted by morphological and hydrological degradations. Correspondingly, fish assemblages seem more structured by local reach-scale pressures than catchment and riparian land use reflecting mostly global water quality problems, while human pressures at these two scales seem to be equally important in structuring macroinvertebrate assemblages.

These results are not in total agreement with the one-out-all-out rule assuming that each BQE responds specifically to a type of pressure. It appears that although BQE responses to human pressures contrast in sensitivity and intensity, most of the BQE indicators do detect the different types of pressure. It is more likely that due to their different sensitivity some BQEs will detect the first impacts of degradations, whereas others will detect severe impacts. According to these results, it is likely that the different biological groups contribute complementary information to the level of degradation of river ecosystems and that **monitoring several groups jointly strengthens the quality of river bio-assessment**.

Observing the local river biota, which human impacts are detectable?

Partial redundancy analysis of the fish and macroinvertebrate communities (***P2***) and the comparison of the responses of the four BQEs to human pressures (fish, macroinvertebrate, diatom and macrophyte; ***P1***) show that human disturbances act hierarchically on river communities and consequently on bio-indicators (***P3***). Catchment and riparian forested areas appear to be useful surrogates of the global water quality of the upstream river. However, local variables are indispensable to describe hydro-morphological and local water quality degradations of river ecosystems (***P2***). In agreement with previous studies (e.g. Hering et al. 2006), the combination of different types of human disturbance seems to have the strongest impacts on river biota (***P1***, ***P2***, ***P3***). Among the individual types of pressure, the presence of an impoundment is known to modify the functioning of the river totally (Baxter 1977); for this reason, it was considered separately. With water quality degradation, it appears to affect river biota the most (***P1***, ***P2***).

This hierarchy assumes that river biota are generally very sensitive to water quality and therefore that improving water quality may substantially improve the ecological health of the river. By contrast, improving the river's hydro-morphological conditions without dealing with water quality problems may not enable the improvement of a river's ecological "health". This result emphasizes that sites should not be impaired by other types of pressure to study the effects of a particular pressure. Otherwise, the effects are likely to be confounded. However, rivers are often impacted by multiple pressures and very large data sets are required to be able to analyse these individual pressures.

Despite this average common pattern, responses to human pressures were highly variable among metrics and indicators. For instance, the IPR+ index detects the low pressure of water quality and morphological degradations and only medium levels of hydrological degradations. The intensities of the responses to the three individual types of degradation were similar.

Interaction effects between natural and human factors on one hand and different types of human factors on the other

The results of ***P1*** and ***P2*** suggest the importance of complex shared effects between environmental factors and human pressures and among human pressures. Moreover, the fact that biological communities and related metrics are globally more sensitive to a combination of different types of river degradations than to specific degradations raises new issues about the relationship among pressures. Do the different types of pressure have additive, multiplicative or opposite effects? If, as indicated by these preliminary results, interaction effects among pressures exist, in the common case of multi-impacted sites, it will be very hard to separate the effects of single pressures and determine a simple pressure–impact diagnosis. Consequently, it will be difficult to answer the water managers regarding the main pressure disturbing a river's ecological status. Unfortunately, the understanding of the combined or cumulated effect of several types of pressure on river aquatic assemblages is typically poor (Pont et al. 2007).

In addition, if interaction effects exist between human pressure and natural environment features, the responses of aquatic assemblages to human pressures will differ depending on the physiography. In this case, the predictive model approach to discard the effect of the environment will be insufficient because the interaction effects will remain after standardization. If it exists, this unconsidered part of community variability will greatly complicate analysis of the responses to human pressures.

Perspectives

Future advanced research on the effects of complex interactions of human pressures and between pressures and physiography would advance both management and our understanding of river ecosystems.

Testing such interaction effects must be based on a large data set recreating a balanced experimental design. This aspect could be primarily addressed through particular cases of pressures for which effects are well known and processes are at least partly interpretable. For instance, testing the interaction effects between the presence of an impoundment (dam or weir) and the effect of water quality on biota), the main question will be: Are the effects of impoundment and water quality degradation additive or multiplicative? In other words, are the effects of water quality degradation the same for a river impounded by a dam and for a free-flowing river?

The next step: what is the response of bio-indicators to improvement of river conditions?

The literature review of the effects of the removal of small dams on river communities (***P5***) emphasizes the need for long-term monitoring to improve the understanding of the effects of restoration programmes. Whereas some positive effects are expected from restoration measures, it could take years to observe these improvements. By reviewing and structuring existing knowledge, conceptual models can provide a useful framework for planning more effective river improvement measures and for identifying leads for future research in restoration ecology. However, quantitative measures of the effectiveness and success of the recovery are lacking.

Perspectives

Logically, a further step in bio-assessment field development is to measure the response of river ecosystem health to the reduction of human pressures. ***P1***, ***P2*** and ***P3*** show that the biological indicators developed respond to human-induced degradations. The subsequent question is can they measure and predict the degree of success of the measures undertaken. Yet, some doubt still reigns on the capacity of the river ecosystem to return to the reference status. Indeed, it is unlikely that restoration effects are simply the opposite of degradation effects (Moerke et al. 2004; Feld et al. 2011).

As illustrated by weir removal case studies, whereas the water quality and hydrology of the river could recover relatively rapidly, it is likely that the morphological characteristics of the river will take years to recover, or may never return to former conditions and consequently biota will not be improved as much as expected.

This refers to the resilience of river ecosystems and their capacity to return to reference conditions when eliminating the degradation causes. Failures of restoration efforts might stem from changes that have occurred over the past few decades that are not considered in the implementation of the measures such as global changes in land use, general loss of connectivity between ecosystems, loss of native species pools or invasion by non-native species (Suding et al. 2004). Congruently, a growing idea is that human impacts should be managed at the catchment scale (Verdonschot 2000). For instance, water quality problems cannot be resolved at the small scale if pollutants originate from farming, domestic sewage and industrial waste discharges from the whole upstream catchment.

In cases where river ecosystems are able to return to the same functioning as in reference conditions, contrary to compositional indicators, functional bio-indicators will probably detect these improvements. In future research, the extent to which biological indicators detect improvements in river conditions (intensity and sensitivity of the responses) would have to be investigated. According to ***P1*** and ***P2*** results, the less sensitive but more intense responses of fish metrics would be better adapted to the first results of improvement measures, while diatom and macroinvertebrate metrics would be more useful in detecting more advanced stages of restoration. In addition, one can wonder whether the same hierarchy and relationships observed for impacts of human pressures will be observed for the effects of river condition improvement.

On the other hand, in the case where river restoration results in a shift leading to ecosystem functions different from the reference condition, even functional bio-indicators will not be suitable to measure the success of the restoration. This dynamic view of ecosystems challenges the meaning of the definition of a unique reference condition range for each river type (environmental abiotic factors). Another point of view is that setting clear and achievable goals should focus on the desired characteristics for the river in the future, rather than in relation to what they were in the past (Hobbs & Harris 2001; Jähnig et al. 2011).

REFERENCES

AFNOR (2004) Qualité de l'eau - Détermination de l'indice poissons rivières (IPR). Association française de normalisation. Norme homologuée T90-344.

AFNOR (2007) Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD). Association française de normalisation. Norme homologuée NF T90-354.

Allan, J.D. & Castillo, M.M. (2007) Stream ecology: Structure and function of running waters. – Kluwer Academic Publishers.

Archaimbault, V., Usseglio-Polatera, P., Garric, J., Wasson, J.-G. & Babut, M. (2010) Assessing pollution of toxic sediment in streams using bio-ecological traits of benthic macroinvertebrates. Freshwater Biology, 55(7), 1430-46.

Baattrup-Pedersen, A. & Riis, T. (1999) Macrophyte diversity and composition in relation to substratum characteristics in regulated and unregulated Danish streams. Freshwater Biology, 42(2), 375-85.

Bady, P., Doledec, S., Fesl, C., Gayraud, S., Bacchi, M. & Scholl, F. (2005) Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. Freshwater Biology, 50(1), 159-73.

Bailey, R.C., Kennedy, M.G., Dervish, M.Z. & Taylor R.M. (1998) Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrates communities in Yukon streams. Freshwater Biology, 39, 765-774.

Baron, J.S., Poff, N.L., Angermeier, P.L., Dahm, C.N., Gleick, P.H., Hairston, N.G., Jackson, R.B., Johnston, C.A., Richter, B.D. & Steinman, A.D. (2002) MEETING ECOLOGICAL AND SOCIETAL NEEDS FOR FRESHWATER. Ecological Applications, 12(5), 1247-60.

Baxter, R.M. (1977) Environmental Effects of Dams and Impoundments. Annual Review of Ecology and Systematics, 8, 255-83.

Bêche, L.A., McElravy, E.P. & Resh, V.H. (2006) Long-term seasonal variation in the biological traits of benthic-macroinvertebrates in two Mediterranean-climate streams in California, U.S.A. Freshwater Biology, 51(1), 56-75.

Bednarek, A.T. (2001) Undamming rivers: A review of the ecological impacts of dam removal. Environmental Management, 27(6), 803-14.

Bernhardt, E.S., Palmer, M.A., Allan, J.D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G.M., Lake, P.S., Lave, R., Meyer, J.L., O'Donnell, T.K., Pagano, L., Powell, B. & Sudduth, E. (2005) Synthesizing U.S. River Restoration Efforts. Science, 308(5722), 636-37.

Besse-Lototskaya, A., Verdonschot, P.F.M., Coste, M. & Van de Vijver, B. (2011) Evaluation of European diatom trophic indices. Ecological Indicators, 11(2), 456-67.

Bis, B., Zdanowicz, A. & Zalewski, M. (2000) Effects of catchment properties on hydrochemistry, habitat complexity and invertebrate community structure in a lowland river. Hydrobiologia, 422-423, 369-87.

Bond, N.R. & Lake, P.S. (2003) Characterizing fish-habitat associations in streams as the first step in ecological restoration. Austral Ecology, 28(6), 611-21.

Borcard, D., Gillet F. & Legendre, P. (2011) Numerical ecology with R. Springer, Use R! series, New York.

Born, S.M., Filbert, T.L., Genskow, K.D., Hernandez-Mora, N., Keefer, M.L. & White, K.A. (1996). The removal of small dams: An institutional analysis of the Wisconsin experience. Cooperative Extension Report 96-1, May. Department of Urban and Regional Planning, University of Wisconsin–Madison.

Born, S.M., Genskow, K.D., Filbert, T.L., Hernandez-Mora, N., Keefer, M.L. & White, K.A. (1998) Socioeconomic and institutional dimensions of dam removals: The Wisconsin experience. Environmental Management, 22(3), 359-70.

Boston, J., Martin, J., Pallot, J. & Walsh, P. (1996) Public Management: The New Zealand Model, Oxford University Press, UK.

Bouleau, G. & Pont, D. (submitted) The reference conditions in the Water framework directive facing the dynamics of hydrosystems and their uses.

Brooks, S.P. (2003) Bayesian computation: a statistical revolution. Philosophical Transactions of the Royal Society of London. Series A:

Bunn, S.E. & Arthington, A.H. (2002) Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management, 30(4), 492-507.

Bunn, S.E. & Davies, P.M. (2000) Biological processes in running waters and their implications for the assessment of ecological integrity. Hydrobiologia, 422-423, 61-70.

Bushaw-Newton, K.L., Hart, D.D., Pizzuto, J.E., Thomson, J.R., Egan, J., Ashley, J.T., Johnson, T.E., Horwitz, R.J., Keeley, M., Lawrence, J., Charles, D., Gatenby, C., Kreeger, D.A., Nightengale, T., Thomas, R.L. & Velinsky, D.J. (2002) An integrative approach towards understanding ecological responses to dam removal: The Manatawny Creek Study. Journal of the American Water Resources Association, 38(6), 1581-99.

Carstensen, J. (2007) Statistical principles for ecological status classification of Water Framework Directive monitoring data. Marine Pollution Bulletin, 55(1-6), 3-15.

Chaplin, J.J. (2003) Framework for monitoring and preliminary results after removal of Good Hope Mill Dam. In: W.L. Graf, Editor, Dam Removal Research: Status and Prospects, The H. John Heinz III Center for Science, Economics and the Environment, Washington, DC.

Cheng, F. & Granata, T. (2007) Sediment transport and channel adjustments associated with dam removal: Field observations. Water Resources Research, 43(3).

Chessman, B.C., Fryirs, K.A. & Brierley, G.J. (2006) Linking geomorphic character, behaviour and condition to fluvial biodiversity: implications for river management. Aquatic Conservation-Marine and Freshwater Ecosystems, 16(3), 267-88.

CIS (2012) Ecostat Hydromorphology, Workshop in Brussels on 12-13 June 2012.

Clarke, R.T. & Hering, D. (2006) Errors and uncertainty in bioassessment methods - Major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia, 566(1), 433-39.

Clarke, R.T., Furse, M.T., Wright, J.F. & Moss, D. (1996) Derivation of a biological quality index for river sites: Comparison of the observed with the expected fauna. Journal of Applied Statistics, 23(2-3), 311-32.

Clarke, R.T., Wright, J.F. & Furse, M.T. (2003) RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modelling, 160(3), 219-33.

Collier, K.J. (2008) Temporal patterns in the stability, persistence and condition of stream macroinvertebrate communities: relationships with catchment land-use and regional climate. Freshwater Biology, 53(3), 603-16.

Daan, N. (2005) An Afterthought: Ecosystem Metrics and Pressure Indicators. ICES Journal of Marine Science: Journal du Conseil, 62(3), 612-13.

Dolédec, S., Phillips, N., Scarsbrook, M., Riley, R.H. & Townsend, C.R. (2006) Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. Journal of the North American Benthological Society, 25(1), 44-60.

Doyle, M.W., Stanley, E.H., Orr, C.H., Selle, A.R., Sethi, S.A. & Harbor, J.M. (2005) Stream ecosystem response to small dam removal: Lessons from the Heartland. Geomorphology, 71(1-2), 227-44.

Durance, I., Lepichon, C. & Ormerod, S.J. (2006) Recognizing the importance of scale in the ecology and management of riverine fish. River Research and Applications, 22(10), 1143-52.

EEA (1995) Inland Waters - Europe's Environment: The Dobris Assessment 1994 (chapter 5), European Environmental Agency.

Ellis, J.E. (2007) Combining multiple quality elements and defining spatial rules for WFD classification. Environment Agency Science Project Number SC060044, Product Code GEHO0807BMXZ-E-E.

European Union (2000) Directive 2000/60/EC of the European Parliament and of the council establishing a framework for the community action in the field of water policy. Off. J. Eur. Commun. L327: 1-72.

Fausch, K.D., Torgersen, C.E., Baxter, C.V. & Li, H.W. (2002) Landscapes to riverscapes: Bridging the gap between research and conservation of stream fishes. Bioscience, 52(6), 483-98.

Feld, C.K., Birk, S., Bradley, D.C., Hering, D., Kail, J., Marzin, A., Melcher, A., Nemitz, D., Pedersen, M.L., Pletterbauer, F., Pont, D., Verdonschot, P.F.M., Friberg, N. & Guy, W. (2011). From Natural to Degraded Rivers and Back Again: A Test of Restoration Ecology Theory and Practice. In Advances in Ecological Research, Vol. Volume 44, pp. 119-209. Academic Press.

Ferréol, M., Dohet, A., Cauchie, H.-M. & Hoffmann, L. (2005) A Top-down Approach for the Development of a Stream Typology Based on Abiotic Variables. Hydrobiologia, 551(1), 193-208.

Finetti, B. (1970) Logical foundations and measurement of subjective probability. Acta Psychologica, 34(0), 129-45.

Fore, L.S., Karr, J.R. & Conquest, L.L. (1994) Statistical Properties of an Index of Biological Integrity Used to Evaluate Water-Resources. Canadian Journal of Fisheries and Aquatic Sciences, 51(5), 1077-87.

Franssen, N.R., Tobler, M. & Gido, K.B. (2011) Annual variation of community biomass is lower in more diverse stream fish communities. Oikos, 120(4), 582-90.

Friberg, N., Bonada, N., Bradley, D.C., Dunbar, M.J., Edwards, F.K., Grey, J., Hayes, R.B., Hildrew, A.G., Lamouroux, N., Trimmer, M. & Woodward, G. (2011) Biomonitoring of human impacts in freshwater ecosystems: the good, the bad, and the ugly. Advances in Ecological Research, 44, 1-68.

Funtowicz, S.O. & Ravetz, J.R. (1993) Science for the post-normal age. Futures, 25(7), 739-55.

Furse, M.T., Hering, D., Brabec, K., Buffagni, A., Sandin, L. & Verdonschot, P.F.M. (2006) The ecological status of European rivers: Evaluation and intercalibration of assessment methods. Hydrobiologia, 566(1), 1-2.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2004) Bayesian data analysis. Chapman and Hall, London.

Gordon, N.D., McMahon, T.A., Finlayson, B.L., Gippel, C.J. & Nathan, R.J. (2004) Stream hydrology. An introduction for ecologists. – John Wiley & Sons Ltd.

Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4(4), 379-91.

Gregory, S., Li, H. & Li, J. (2002) The conceptual basis for ecological responses to dam removal. Bioscience, 52(8), 713-23.

Hacking, I. (1965). The Logic of Statistical Inference. Cambridge University Press.

Hamilton, A.T., Barbour, M.T. & Bierwagen, B.G. (2010) Implications of global change for the maintenance of water quality and ecological integrity in the context of current water laws and environmental policies. Hydrobiologia, 657(1), 263-78.

Hatton-Ellis, T. (2008) The Hitchhiker's Guide to the Water Framework Directive. Aquatic Conservation: Marine and Freshwater Ecosystems, 18(2), 111-16.

Hawkins, C.P., Norris, R.H., Hogue, J.N. & Feminella, J.W. (2000) Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications, 10, 1456-1477.

Heino, J., Paavola, R., Virtanen, R. & Muotka, T. (2005) Searching for biodiversity indicators in running waters: do bryophytes, macroinvertebrates, and fish show congruent diversity patterns? Biodiversity and Conservation, 14(2), 415-28.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C.K., Heiskanen, A.-S., Johnson, R.K., Moe, J., Pont, D., Solheim, A.L. & de Bund, W.v. (2010) The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future. Science of The Total Environment, 408(19), 4007-19.

Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K. & Verdonschot, P.F.M. (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: A comparative metric-based analysis of organism response to stress. Freshwater Biology, 51(9), 1757-85.

Hering, D., Moog, O., Sandin, L. & Verdonschot, P.F.M. (2004) Overview and application of the AQEM assessment system. Hydrobiologia, 516(1-3), 1-20.

Hill, M.J., Long, E.A., and Hardin, S. (1994). Effects of dam removal on Dead Lake, Chipola River, Florida. Apalachi-cola River Watershed Investigations, Florida Game and Fresh Water Fish Commission. A Wallop-Breaux Project F-39-R.

Hobbs, R.J. & Harris, J.A. (2001) Restoration Ecology: Repairing the Earth's Ecosystems in the New Millennium. Restoration Ecology, 9(2), 239-46.

Hoeinghaus, D.J., Winemiller, K.O. & Birnbaum, J.S. (2007) Local and regional determinants of stream fish assemblage structure: Inferences based on taxonomic vs. functional groups. Journal of Biogeography, 34(2), 324-38.

Holt, E.A. & Miller, S.W. (2011) Bioindicators: Using Organisms to Measure Environmental Impacts. Nature Education Knowledge 2(2):8

Horwitz, R.J. (1978) Temporal Variability Patterns and the Distributional Patterns of Stream Fishes. Ecological Monographs, 48(3), 307-21.

Hrodey, P.J., Sutton, T.M., Frimpong, E.A. & Simon, T.P. (2009) Land-use Impacts on Watershed Health and Integrity in Indiana Warmwater Streams. The American Midland Naturalist, 161(1), 76-95.

Huang, G.H. & Chang, N.B. (2003) Perspectives of Environmental Informatics and Systems Analysis. Journal of Environmental Informatics, 1(1), 1-6.

Huet, M. (1954) Biologie, profils en l o ng et e n travers des eaux courantes. Bull. Fr Piscic, 11, 32-351.

Hughes, R.M., Howlin, S. & Kaufmann, P.R. (2004) A biointegrity index (IBI) for coldwater streams of Western Oregon and Washington. *Transactions of the American Fisheries Society,* 133, 1497-1515.

Hughes, S.J., Santos, J.M., Ferreira, M.T., Caraça, R. & Mendes, A.M. (2009) Ecological assessment of an intermittent Mediterranean river using community structure and function: Evaluating the role of different organism groups. Freshwater Biology, 54(11), 2383-400.

Hugueny, B., Camara, S., Samoura, B. & Magassouba, M. (1996) Applying an index of biotic integrity based on fish assemblages in a West African river. Hydrobiologia, 331(1-3), 71-78.

Hutchinson, G.E. (1957) Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22: 415–427.

Hynes, H.B.N. (1960) The biology of polluted waters. Liverpool University Press, Liverpool, UK.

Iversen, T.M., Kronvang, B., Madsen, B.L., Markmann, P. & Nielsen, M.B. (1993) Re-establishment of Danish streams: Restoration and maintenance measures. Aquatic Conservation: Marine and Freshwater Ecosystems, 3(2), 73-92.

Jähnig, S.C., Lorenz, A.W., Hering, D., Antons, C., Sundermann, A., Jedicke, E. & Haase, P. (2011) River restoration success: a question of perception. Ecological Applications, 21, 2007-2015.

Jenkins, T.M., Diehl, S., Kratz, K.W. & Cooper, S.D. (1999) Effects of population density on individual growth of brown trout in streams. Ecology, 80(3), 941-56.

Johnson, R.K. & Hering, D. (2009) Response of taxonomic groups in streams to gradients in resource and habitat characteristics. Journal of Applied Ecology, 46(1), 175-86.

Johnson, R.K., Hering, D., Furse, M.T. & Clarke, R.T. (2006b) Detection of ecological change using multiple organism groups: Metrics and uncertainty. Hydrobiologia, 566(1), 115-37.

Johnson, R.K., Hering, D., Furse, M.T. & Verdonschot, P.F.M. (2006a) Indicators of ecological change: Comparison of the early response of four organism groups to stress gradients. Hydrobiologia, 566(1), 139-52.

Justus, B.G., Petersen, J.C., Femmer, S.R., Davis, J.V. & Wallace, J.E. (2010) A comparison of algal, macroinvertebrate, and fish assemblage indices for assessing low-level nutrient enrichment in wadeable Ozark streams. Ecological Indicators, 10(3), 627-38.

Kanehl, P.D., Lyons, J. & Nelson, J.E. (1997) Changes in the habitat and fish community of the Milwaukee River, Wisconsin, following removal of the Woolen Mills Dam. North American Journal of Fisheries Management, 17(2), 387-400.

Karr, J.R (2006) Seven foundations of biological monitoring and assessment. Biologia Ambientale 20(2): 7-18.

Karr, J.R. & Chu, E.W. (2000) Sustaining living rivers. Hydrobiologia, 422-423, 1-14.

Karr, J.R. (1981) Assessment of Biotic Integrity Using Fish Communities. Fisheries, 6(6), 21-27.

Karr, J.R. (1991) Biological Integrity: A Long-Neglected Aspect of Water Resource Management. Ecological Applications, 1(1), 66-84.

Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., & Schlosser, I.J. (1986) Assessing biological integrity in running waters: a method and its rationale. Ill. Nat. Hist. Surv. Spec. Publ. 5, Urbana, Ill.

Kaspersson, R. & Hojesjo, J. (2009) Density-dependent growth rate in an age-structured population: a field study on stream-dwelling brown trout Salmo trutta. Journal of Fish Biology, 74(10), 2196-215.

Kolkwitz, R. & Marsson, M. (1902) Grundsätze für die biologische Beurtheilung des Wassers nach seiner Flora und Fauna. Mitt. Königl. Prüfungsanstalt Wasser Abwasser 1:3–72

Lacoul, P. & Freedman, B. (2006) Environmental influences on aquatic plants in freshwater ecosystems. Environmental Reviews, 14(2), 89-136.

Lafont, M., Camus, J.C. & Rosso, A. (1996) Superficial and hyporheic oligochaete communities as indicators of pollution and water exchange in the River Moselle, France. Hydrobiologia, 334(1), 147-55.

Lammert, M. & Allan, J.D. (1999) Assessing biotic integrity of streams: Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management, 23(2), 257-70.

Lamouroux, N., Poff, N.L. & Angermeier, P.L. (2002) Intercontinental convergence of stream fish community traits along geomorphic and hydraulic gradients. Ecology, 83(7), 1792-807.

Leaniz, C.G. (2008) Weir removal in salmonid streams: implications, challenges and practicalities. Hydrobiologia, 609, 83-96.

Legendre, P. & Gallagher, E. (2001) Ecologically meaningful transformations for ordination of species data. Oecologia, 129(2), 271-80.

Leopold, L.B., Wolman, M.G. & Miller, J.P. (1992) Fluvial processes in geomorphology. – Dover Publications, Inc. Petts, G. E. and Amoros, C. 1996. Fluvial hydrosystems. Chapman & Hall.

Leunda, P.M., Sistiaga, M., Oscoz, J. & Miranda, R. (2012) Ichthyofauna of a near-Natural Pyrenean River: Spatio-Temporal Variability and Reach-Scale Habitat. Environmental Engineering and Management Journal, 11(6), 1111-24.

Linke, S., Bailey, R.C. & Schwindt, J. (1999) Temporal variability of stream bioassessments using benthic macroinvertebrates. Freshwater Biology, 42(3), 575-84.

Logez, M. & Pont, D. (2012) Global warming and potential shift in reference conditions: the case of functional fish-based metrics. Hydrobiologia, 1-20.

Logez, M., Bady, P., Melcher, A. & Pont, D. (2012) A continental-scale analysis of fish assemblage unctional structure in European rivers. Ecography. doi: 10.1111/j.1600-0587.2012.07447.x

Logez, M., Pont, D. & Ferreira, M.T. (2010) Do Iberian and European fish faunas exhibit convergent functional structure along environmental gradients? Journal of the North American Benthological Society, 29(4), 1310-23.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. & Randers, J. (2005) The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society B-Biological Sciences, 360(1454), 289-95.

Lorenz, A., Feld, C.K. & Hering, D. (2004) Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. Limnologica - Ecology and Management of Inland Waters, 34(4), 379-89.

Manly, B.J.F. (1997) Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman & Hall, London.

Mazor, R.D., Purcell, A.H. & Resh, V.H. (2009) Long-Term Variability in Bioassessments: A Twenty-Year Study from Two Northern California Streams. Environmental Management, 43(6), 1269-86.

McCarthy, M.A. (2007) Bayesian methods for ecology. Cambridge University Press, Cambridge.

McCormick, F.H., Hughes, R.M., Kaufmann, P.R., Peck, D.V., Stoddard, J.L. & Herlihy, A.T. (2001) Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Transactions of the American Fisheries Society, 130(5), 857-77.

McCullagh, P. & Nelder, J.A. (1989) Generalized Linear Models. 2nd edition. Chapman and Hall, London.

Moerke, A.H. & Lamberti, G.A. (2006) Scale-dependent influences on water quality, habitat, and fish communities in streams of the Kalamazoo River Basin, Michigan (USA). Aquatic Sciences, 68(2), 193-205.

Moerke, A.H., Gerard, K.J., Latimore, J.A., Hellenthal, R.A. & Lamberti, G.A. (2004) Restoration of an Indiana, USA, stream: bridging the gap between basic and applied lotic ecology. Journal of the North American Benthological Society, 23, 647-660.

Naiman, R.J. & Latterell, J.J. (2005) Principles for linking fish habitat to fisheries management and conservation. Journal of Fish Biology, 67, 166-85.

Nerbonne, B.A. & Vondracek, B. (2001) Effects of local land use on physical habitat, benthic macroinvertebrates, and fish in the Whitewater River, Minnesota, USA. Environmental Management, 28(1), 87-99.

Nichols, S.J., Robinson, W.A. & Norris, R.H. (2010) Using the reference condition maintains the integrity of a bioassessment program in a changing climate. Journal of the North American Benthological Society, 29, 1459-1471

Oberdorff, T. & Hughes, R.M. (1992) Modification of an index of biotic integrity based on fish assemblages to characterize rivers of the Seine Basin, France. Hydrobiologia, 228(2), 117-30.

Oberdorff, T. & Porcher, J.P. (1992) Fish assemblage structure in Brittany streams (France). Aquat. Living. Resour, 5, 215-223.

Oberdorff, T., Pont, D., Hugueny, B. & Chessel, D. (2001) A probabilistic model characterizing fish assemblages of French rivers: A framework for environmental assessment. Freshwater Biology, 46(3), 399-415.

Oberdorff, T., Pont, D., Hugueny, B. & Porcher, J.P. (2002) Development and validation of a fish-based index for the assessment of 'river health' in France. Freshwater Biology, 47(9), 1720-34.

OECD (1993) Organisation for Economic Co-operation and Development (OECD) core set of indicators for environmental performance reviews. OECD Environment Monographs No. 83, Paris.

OECD (1994) Indicateurs d'environnement : Corps central de l'OCDE, OCDE, Paris.

Ofenbock, T., Moog, O., Gerritsen, J. & Barbour, M. (2004) A stressor specific multimetric approach for monitoring running waters in Austria using benthic macro-invertebrates. Hydrobiologia, 516(1-3), 251-68.

Orr, C.H., Rogers, K.L. & Stanley, E.H. (2006) Channel morphology and P uptake following removal of a small dam. Journal of the North American Benthological Society, 25(3), 556-68.

Ostermiller, J.D. & Hawkins, C.P. (2004) Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society, 23(2), 363-82.

Palmer, M., Allan, J.D., Meyer, J. & Bernhardt, E.S. (2007) River restoration in the twenty-first century: Data and experiential future efforts. Restoration Ecology, 15(3), 472-81.

Palmer, M.A., Bernhardt, E.S., Allan, J.D., Lake, P.S., Alexander, G., Brooks, S., Carr, J., Clayton, S., Dahm, C.N., Shah, J.F., Galat, D.L., Loss, S.G., Goodwin, P., Hart, D.D., Hassett, B., Jenkinson, R., Kondolf, G.M., Lave, R., Meyer, J.L., O'Donnell, T.K., Pagano, L. & Sudduth, E. (2005) Standards for ecologically successful river restoration. Journal of Applied Ecology, 42(2), 208-17.

Pantle, R. & Buck, H. (1955) Die biologische Uberwachung der Gewasser und die Darstellung der Ergebnisse. Besondere Mitteilung im Deutschen Gewasserkundlichen. 12 (135-143)

Pardo, I., Gomez-Rodriguez, C., Wasson, J.G., Owen, R., van de Bund, W., Kelly, M., Bennett, C., Birk, S., Buffagni, A., Erba, S., Mengin, N., Murray-Bligh, J. & Ofenboeck, G. (2012) The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. Science of The Total Environment, 420, 33-42.

Peres-Neto, P.R., Legendre, P., Dray, S. & Borcard, D. (2006) Variation partitioning of species data matrices: Estimation and comparison of fractions. Ecology, 87(10), 2614-25.

Petts, G.E. (1984) Impounded rivers: Perspectives for ecological management. John Wiley & Sons. Chichester, England.

Petts, G.E. (1989) Historical analysis of fluvial hydrosystems, in: Petts, G.E., Muller, H. and Roux, A.L. (Eds.), Historical Change of Large Alluvial Rivers: Western Europe. Wiley, Chichester, pp. 1-18.

Poff, N.L. (1997) Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society, 16(2), 391-409.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E. & Stromberg, J.C. (1997) The Natural Flow Regime. Bioscience, 47, 769-784.

Pollard, A.I. & Reed, T. (2004) Benthic invertebrate assemblage change following dam removal in a Wisconsin stream. Hydrobiologia, 513(1-3), 51-58.

Pont, D., Delaigue, O. & Belliard, J. (2012) Programme IPR+ : révision de l'indice poisson rivière pour l'application de la DCE. Rapport Partenariat 2011 – Mise au point d'un indicateur poisson IPR+ – Action 12.

Pont, D., Hugueny, B. & Oberdorff, T. (2005) Modelling habitat requirement of European fishes: Do species have similar responses to local and regional environmental constraints? Canadian Journal of Fisheries and Aquatic Sciences, 62(1), 163-73.

Pont, D., Hugueny, B. & Rogers, C. (2007) Development of a fish-based index for the assessment of river health in Europe: The European Fish Index. Fisheries Management and Ecology, 14(6), 427-39.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N. & Schmutz, S. (2006) Assessing river biotic condition at a continental scale: A European approach using functional metrics and fish assemblages. Journal of Applied Ecology, 43(1), 70-80.

Punt, A. & Hilborn, R.A.Y. (1997) Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7(1), 35-63.

Randle, T.J. (2003). Dam removal and sediment management. In: Dam Removal Research: Status and Prospects (Ed. by T.J. Randle and W.L. Graf), pp. 81–104. The Heinz Center for Science, Economics, and the Environment, Washington, DC.

Richards, C., Haro, R.J., Johnson, L.B. & Host, G.E. (1997) Catchment and reach-scale properties as indicators of macroinvertebrate species traits. Freshwater Biology, 37(1), 219-30.

Ricklefs, R.E. & Schluter, D. (1993). Species diversity in ecological communities: historical and geographical perspectives. In Species diversity in ecological communities: historical and geographical perspectives., p 414 pp. University of Chicago Press.

Ricklefs, R.E. (1987) Community Diversity: Relative Roles of Local and Regional Processes. Science, 235(4785), 167-71.

Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. Ecology Letters, 7(1), 1-15.

Rolauffs, P., Stubauer, I., Zahradkova, S., Brabec, K. & Moog, O. (2004) Integration of the saprobic system into the European Union Water Framework Directive - Case studies in Austria, Germany and Czech Republic. Hydrobiologia, 516(1), 285-98.

Ross, S.T., Matthews, W.J. & Echelle, A.A. (1985) Persistence of Stream Fish Assemblages - Effects of Environmental-Change. American Naturalist, 126(1), 24-40.

Saly, P., Takacs, P., Kiss, I., Biro, P. & Eros, T. (2011) The relative influence of spatial context and catchment- and site-scale environmental factors on stream fish assemblages in a human-modified landscape. Ecology of Freshwater Fish, 20(2), 251-62.

Sandin, L. & Verdonschot, P.F.M. (2006) Stream and river typologies - Major results and conclusions from the STAR project. Hydrobiologia, 566, 33-37.

Schaefer, J.F., Clark, S.R. & Warren, M.L. (2012) Diversity and stability in Mississippi stream fish assemblages. Freshwater Science, 31(3), 882-94.

Schlosser, I.J. (1982) Fish Community Structure and Function Along 2 Habitat Gradients in a Headwater Stream. Ecological Monographs, 52(4), 395-414.

Schmutz, S., Cowx, I.G., Haidvogl, G. & Pont, D. (2007) Fish-based methods for assessing European running waters: A synthesis. Fisheries Management and Ecology, 14(6), 369-80.

Shafroth, P.B., Friedman, J.M., Auble, G.T., Scott, M.L. & Braatne, J.H. (2002) Potential Responses of Riparian Vegetation to Dam Removal. BioScience, 52(8), 703-12.

Shaw, W.D. & Woodward, R.T. (2010) Water Management, Risk, and Uncertainty: Things We Wish We Knew in the 21st Century. Western Economic Forum, 9.

Shuman, J.R. (1995) Environmental considerations for assessing dam removal alternatives for river restoration. Regulated Rivers-Research & Management, 11(3-4), 249-61.

Sládeček, V. (1965) The future of the saprobity system. Hydrobiologia, 25(3-4), 518-37.

Snyder, C.D., Young, J.A., Villella, R. & Lemarie, D.P. (2003) Influences of upland and riparian land use patterns on stream biotic integrity. Landscape Ecology, 18(7), 647-64.

Sousa, W.P. (1984) The Role of Disturbance in Natural Communities. Annual Review of Ecology and Systematics, 15(1), 353-91.

Southwood, T.R.E. (1977) Habitat, Templet for Ecological Strategies - Presidential-Address to British-Ecological-Society, 5 January 1977. Journal of Animal Ecology, 46(2), 337-65.

Spiegelhalter, D.T.A., Best, N., & Lunn, D. (2003) WinBUGS user manual. Version 1.4. Available at http://www.mrc-bsu.cam.ac.uk/bugs.

Stanford, J.A., Ward, J.V., Liss, W.J., Frissell, C.A., Williams, R.N., Lichatowich, J.A. & Coutant, C.C. (1996) A general protocol for restoration of regulated rivers. Regulated Rivers: Research and Management, 12, 391-413.

Staniszewski, R., Szoszkiewicz, K., Zbierska, J., Lesny, J., Jusik, S. & Clarke, R.T. (2006) Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification. Hydrobiologia, 566, 235-46.

Stanley, E.H. & Doyle, M.W. (2003) Trading off: the ecological removal effects of dam removal. Frontiers in Ecology and the Environment, 1(1), 15-22.

Statzner, B., Bis, B., Dolédec, S. & Usseglio-Polatera, P. (2001) Perspectives for biomonitoring at large spatial scales: A unified measure for the functional composition of invertebrate communities in European running waters. Basic and Applied Ecology, 2(1), 73-85.

Stearns, S.C. (1992) The Evolution of Life Histories. Oxford University Press, Oxford, New York.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K. & Norris, R.H. (2006) Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications, 16(4), 1267-76.

Suding, K.N., Gross, K.L. & Houseman, G.R. (2004) Alternative states and positive feedbacks in restoration ecology, Trends in Ecology and Evolution, 19, 1, 46-53.

Tachet, H., Richoux, P., Bournaud, M. & Ussegio-Polatera, P. (2006) Invertébrés d'eau douce: systématique, biologie, écologie. CNRS Editions, Paris.

Thompson, R.M., Dunne, J.A. & Woodward, G. (2012) Freshwater food webs: towards a more fundamental understanding of biodiversity and community dynamics. Freshwater Biology, 57, 1329-1341.

Thomson, J.R., Hart, D.D., Charles, D.F., Nightengale, T.L. & Winter, D.M. (2005) Effects of removal of a small dam on downstream macroinvertebrate and algal assemblages in a Pennsylvania stream. Journal of the North American Benthological Society, 24(1), 192-207.

Tiemann, J.S., Gillette, D.P., Wildhaber, M.L. & Edds, D.R. (2004) Effects of lowhead dams on riffle-dwelling fishes and macroinvertebrates in a midwestern river. Transactions of the American Fisheries Society, 133(3), 705-17.

Todd, C.R. & Burgman, M.A. (1998) Assessment of Threat and Conservation Priorities under Realistic Levels of Uncertainty and Reliability. Conservation Biology, 12(5), 966-74.

Tomanova, S., Tedesco, P.A., Roset, N., Berrebi dit Thomas, R. & Belliard, J. (in revision in Fisheries Management and Ecology) Systematic point sampling of fish communities in medium and large-sized rivers: sampling procedure and effort.

Tonn, W.M., Magnuson, J.J., Rask, M. & Toivonen, J. (1990) Intercontinental comparison of small-lake fish assemblages: the balance between local and regional processes. American Naturalist, 136(3), 345-75.

Townsend, C.R. & Hildrew, A.G. (1994) Species Traits in Relation to a Habitat Templet for River Systems. Freshwater Biology, 31(3), 265-75.

Usseglio-Polatera, P., Bournaud, M., Richoux, P. & Tachet, H. (2000) Biomonitoring through biological traits of benthic macroinvertebrates: How to use species trait databases? Hydrobiologia, 422-423, 153-62.

van der Sluijs, J.P. (2007) Uncertainty and precaution in environmental management: Insights from the UPEM conference. Environmental Modelling and Software, 22(5), 590-98.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. (1980) The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences, 37(1), 130-37.

Verdonschot, P. & Nijboer, R. (2004) Testing the European stream typology of the Water Framework Directive for macroinvertebrates. Hydrobiologia, 516(1), 35-54.

Verdonschot, P. (2006) Beyond Masses and Blooms: The Indicative Value of Oligochaetes. Hydrobiologia, 564(1), 127-42.

Verdonschot, P.F.M. (2000) Integrated ecological assessment methods as a basis for sustainable catchment management. Hydrobiologia, 422-423, 389-412.

Wang, L.Z., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. & Stewart, P.M. (2003) Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences, 60(5), 491-505.

Ward, J.V. & Stanford, J.A. (1979) The Ecology of Regulated Streams. Plenum Press, New York, NY.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. Annual Review of Ecology and Systematics, 33(1), 475-505.

Willby N., Birk S. & Bonne W. (2010) IC Guidance Annex V: Definition of comparability criteria for setting class boundaries, 22 p.

Wolda, H. (1981) Similarity Indexes, Sample-Size and Diversity. Oecologia, 50(3), 296-302.

Working Group Ecostat (2009) Implementation strategy for the Water Framework Directive (2000/60/Ec) Guidance document no. 14. Guidance document on the intercalibration process 2008-2011, 55 p.

Wright, J.F., Armitage, P.D., Furse, M.T. & Moss, D. (1993) RIVPACs – a technique for evaluating the biological quality of rivers in the UK. European Water Pollution Control, 3, 15–25.

Yates, A.G. & Bailey, R.C. (2010) Covarying patterns of macroinvertebrate and fish assemblages along natural and human activity gradients: Implications for bioassessment. Hydrobiologia, 637, 87-100.

Zelinka, M. & Marvan, P. (1961) Zur Prazisierung der biologischen Klassifikation der Reinheit fliessender Gewasser. Arch. Hydrobiol. 57 (3): 389-407.

PART II: JOURNAL ARTICLES

P1

Ecological assessment of running waters: Do macrophytes, macroinvertebrates, diatoms and fish show similar responses to human pressures?

Marzin, A., V. Archaimbault, J. Belliard, C. Chauvin, F. Delmas & D. Pont (2012)

Ecological Indicators 23: 56-65.

Ecological Indicators 23 (2012) 56–65

Contents lists available at SciVerse ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Ecological assessment of running waters: Do macrophytes, macroinvertebrates, diatoms and fish show similar responses to human pressures?

Anahita Marzin[a,*], Virginie Archaimbault[a], Jérôme Belliard[a], Christian Chauvin[b], François Delmas[b], Didier Pont[a]

[a] Irstea, UR HBAN, HYNES (Irstea – EDF R&D), 1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony Cedex, France
[b] Irstea, UR REBX, 50 avenue de Verdun, 33612 Cestas Cedex, France

ARTICLE INFO

Article history:
Received 13 September 2011
Received in revised form 7 March 2012
Accepted 13 March 2012

Keywords:
Macroinvertebrates
Fish
Diatoms
Macrophytes
Bioindicators
River assessment

ABSTRACT

This study aimed to compare the intensity and the sensitivity of the responses of four river biological quality elements (BQEs) – macrophytes, fish, diatoms and macroinvertebrates – to human pressures excluding natural variations in stream ecosystem functioning. Biological, water quality and hydro-morphological data were compiled for 290 French river sites.

Out of the 93 metrics tested, 51 covering the four BQEs responded significantly to global degradation. The responses to specific pressures were consistent with the BQEs' ecological and biological characteristics. For the four BQEs, metrics responded strongly to water quality degradations. Like fish, macroinvertebrate metrics were very sensitive to morphological degradations such as the presence of an impoundment, while diatom and macrophyte metrics did not show strong responses to these changes. Among the four BQEs' metrics, fish metrics responded the strongest to hydrological perturbations. Although a high proportion of the metrics responded only to high levels of human-induced degradations, trait-based metrics seemed the most sensitive and responded to lower levels of pressure. Global and water quality degradations of the river appear to be better detected by BQE metrics than channel morphological and hydrological degradations. Our results highlight the different impacts of human-induced pressures on the four BQE metrics and the challenging task of assessing the effect of single pressures when most of sites are multi-impacted.

1. Introduction

Throughout Europe, streams have experienced a long history of modification by humans (Petts, 1989) and have become one of the most threatened ecosystems (Loh et al., 2005). Since 2000, the Water Framework Directive (EC, 2000) recommends the use of multiple biological quality elements (BQEs) to assess "ecological status" of rivers. Given these institutional needs, freshwater scientists have developed numerous tools based on various concepts and biological indicators. Historically, biological responses were examined through metrics focusing on sensitive taxa (e.g. the Saprobic index, Pantle and Buck, 1955). More recently, Southwood (1977) and Townsend and Hildrew (1994) have put forward the idea that combinations of functional traits (ecological and biological) are selected by habitat conditions through the survival ability of individual organisms relative to others (i.e. their fitness). Such integrative approaches were based on the functional structures of fish (Fausch et al., 1990; Index of Biotic Integrity, IBI: Karr, 1981) and macroinvertebrate assemblages (Statzner et al., 2001; Usseglio-Polatera et al., 2000). It is of primary importance to gain a comparative idea of the sensitivity and efficiency of these different indicators in detecting river human-induced degradations.

Responses to human-induced disturbances in rivers have frequently been analysed separately for macroinvertebrates (Archaimbault et al., 2010; Lorenz et al., 2004; Statzner et al., 2001), diatoms (Besse-Lototskaya et al., 2011; Carpenter and Waite, 2000; Fore and Grafe, 2002), macrophytes (Lacoul and Freedman, 2006; Riis et al., 2000) and fish (D'Ambrosio et al., 2009; Pont et al., 2006, 2007; Yates and Bailey, 2010). Nonetheless, only a few authors have compared different assemblage responses to anthropogenic pressures. These studies have shown that metric response patterns

Abbreviations: BQE, biological quality element; MetIND, indexes; MetFUNC, functional trait-based metrics; MetTAX, metrics based on the taxonomic composition; PG, physiographic gradient; DE, discriminatory efficiency; WS, weighted mean of sensitivity.

* Corresponding author at: Irstea, UR HBAN, 1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony Cedex, France. Tel.: +33 01 40 96 61 21; fax: +33 01 40 96 61 99.
E-mail address: anahita.marzin@irstea.fr (A. Marzin).

1470-160X/$ – see front matter
doi:10.1016/j.ecolind.2012.03.010

(Johnson and Hering, 2009) and robustness (Johnson et al., 2006a) differ considerably among BQEs and stressors and with stream type (Heino, 2010; Hering et al., 2006). For instance, it appears likely that hydro-morphological degradations affect fish and macrophyte assemblages more than diatoms and macroinvertebrates (Hughes et al., 2009; Johnson et al., 2006b). More generally, the responses of the four BQEs seem stronger for water quality than for hydro-morphological degradations (Hering et al., 2006).

Some authors have demonstrated that trait-based metrics (MetFUNC) such as the number of euryplastic taxa show the highest sensitivity to human disturbance (Dolédec et al., 2006; Usseglio-Polatera et al., 2000). However, previous studies often investigated metrics based on the taxonomic composition (MetTAX) such as the total number of species (Heino et al., 2005; Johnson and Hering, 2009) rather than trait-based metrics (Hering et al., 2006; Hughes et al., 2009; Johnson et al., 2006a, 2006b; Justus et al., 2010). In addition, as streams are frequently impacted by multiple stressors, single effects of stressors on BQEs have rarely been assessed (Hughes et al., 2009). Based on this literature review, we expected that biological assemblages would present different responses to human disturbances in terms of intensity (i.e. discriminatory efficiency; Ofenböck et al., 2004) and sensitivity (i.e. impact of a low level of pressure). Also, it was assumed that responses to pressures would be stronger for indexes (MetIND) and MetFUNC than for MetTAX and that the standardization method would allow analysing BQE responses along the main environmental gradients.

Comparing the responses of macrophytes, fish, diatoms and macroinvertebrates to different human pressures, this paper attempts to answer the following questions:

(1) Are the intensity and sensitivity of the responses to a general degradation gradient similar among BQEs?
(2) Do all BQEs detect specific pressures similarly (hydrological, morphological, and water quality degradations)?
(3) Do these responses to specific pressures change when only sites impacted by this pressure are considered?
(4) Which type of metric (MetIND, MetTAX and MetFUNC) is more appropriate to detect human pressure impacts?

We focused on a French data set covering a large range of environmental conditions and human-induced pressures. As pointed out by Stoddard et al. (2006), "natural variability in indicators always occurs" and has to be considered when measuring the deviation of ecosystems from a reference status. In this paper, the reference status was recognized as the minimally disturbed conditions (Stoddard et al., 2006). In contrast to the previous studies cited above, physiography (i.e. environmental factors assumed to be independent of human activity such as the geology or the altitude) effects were beforehand differentiated from human pressure effects, standardizing the metrics (Oberdorff et al., 2001; Pont et al., 2006, 2007).

2. Material and methods

2.1. Data compilation

For each of the four BQEs, one sample from French monitoring programs was compiled for 290 French river sites (Fig. 1) distributed along the main environmental gradients (e.g. altitude, geology). Very large rivers (upstream drainage area > 14,000 km^2) were not considered. When several samples were available, the most recent was chosen (samples from 2005 to 2008). During these programs, fish (AFNOR, 2004a), macroinvertebrates (AFNOR, 2004b), macrophytes (mainly aquatic phanerogams, bryophytes and colonial algae; AFNOR, 2003) and benthic diatoms (AFNOR,

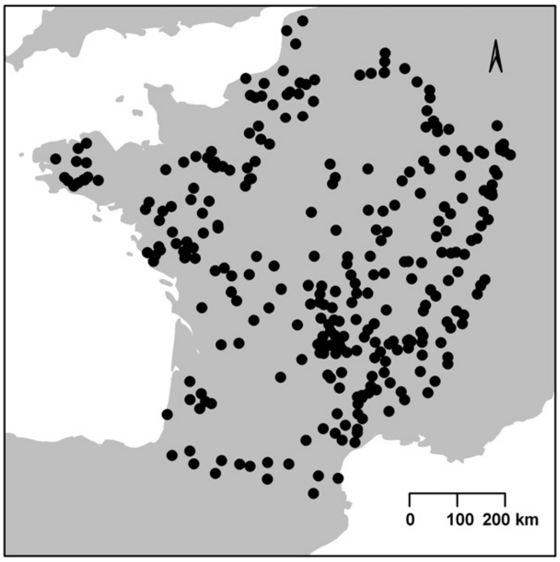


Fig. 1. Location of the 290 sites.

2007) were sampled using standard protocols. Fish, macrophytes and diatoms were identified mostly at the species level and macroinvertebrates at the genus level.

In addition, information on the physiography and human pressures impacting rivers was compiled for each site (Table 1). Pressure data included information on hydro-morphological degradations (e.g. modification of the channel form) and water quality variables (median values for the 2005–2008 period; e.g. nitrate concentration). Most of the rivers ranged from small to medium size, were part of small to relatively large catchments (from 1 to 13,312 km^2; median = 99 km^2) and were situated from 2 to 1520 m above sea level (median = 217 m). Out of the 290 sites, 102 were slightly perturbed, with no or slight hydro-morphological disturbances and water parameters corresponding to a "very good" or "good" status (French stream water quality evaluation system; French Water Agency, 2000). The other 188 sites were considered to be impacted sites. For this study, the 290 sites were divided into two data sets: the calibration data set CAL-80 containing 80 sites randomly selected from the 102 quasi-undisturbed sites and the analysis data set AN-210 including the 188 impacted sites and the 22 other minimally disturbed sites.

2.2. Statistical analysis

2.2.1. Definition of physiographic and human pressure gradients

First, a multivariate analysis was used to summarize the nine physiographic variables (Table 1) into three independent gradients: the first three axes of the analysis PG1, PG2 and PG3. Second, four synthetic human pressure gradients were developed. The global synthetic degradation gradient summarized all the human-induced pressure variables while the three other synthetic gradients were related to human-induced hydrological, morphological and water quality degradations, respectively (Table 1). For each pressure gradient, the first axis of the multivariate analysis was retained and divided into four classes corresponding to four levels of pressure (gp1 = slight pressure to gp4 = strong pressure) using the k-means algorithm (Hartigan and Wong, 1979). It was verified for each variable that the level of pressure increased along the synthetic gradients. PCA (principal component analysis) and MCA (multiple correspondence analysis) were used to analyse quantitative and qualitative variables, respectively. Hill and Smith's analysis (Hill and Smith, 1976) was used to examine quantitative and

Table 1
Physiography, pressure variables and synthetic gradients.

Variables	Transformation	Modalities for qualitative variables (number of sites) and ranges for quantitative variables (median (min-max))
Physiography		
Altitude (m)[a]	log(*x*)	217 (2–1520)
Mean width (m)[a]	log(*x*)	7 (0.5–93)
Mean slope (‰)[a]	log(*x*)	4 (0.1–82)
Catchment area (km^2)[a]	log(*x*)	99 (1–13,312)
Annual mean air temperature (°C)[a]	log(*x*)	10.5 (5–15.5)
Distance to the source (m)[a]	log(*x*)	17 (0.6–372)
Geological type[a]		Siliceous (131)/calcareous (159)
Upstream lakes[a]		No (288)/Yes (2)
Ecoregions[a]		Alps (6)/central highlands (2)/mediterranean (2)/pyrenees (3)/Western highlands (123)/Western plains (154)
Reach scale human modifications		
Presence of an impoundment at the station[b]		No (263)/yes (27)
Hydrological regime modified[c,b]		No (183)/slight (59)/intermediate (26)/high (22)
Hydropeaking[c,b]		No (264)/yes (26)
Water abstraction[c,b]		No (182)/slight (77)/intermediate (10)/high (21)
Riparian vegetation modified[d,b]		No (178)/slight (81)/intermediate (19)/high (12)
Artificial embankment[d,b]		No (253)/partial (22)/yes (15)
Instream habitat modified[d,b]		No (233)/intermediate (34)/high (23)
Channel form modified[d,b]		No (238)/intermediate (31)/high (21)
Cross-section modified[d,b]		No (241)/intermediate (23)/high (26)
Diked[d,b]		No (260)/intermediate (20)/high (10)
Sedimentation[d,b]		No (170)/slight (64)/intermediate (40)/high (16)
Oxygen saturation (%)[e,b]	log(*x*)	94.3 (48.5–112.3)
BOD5 (mg O_2/l)[e,b]	log(*x*)	1.4 (0.5–4.5)
Nitrite (mg NO_2/l)[e,b]	log(*x*)	0.03 (0.01–0.34)
Nitrate (mg NO_3/l)[e,b]	log(*x*)	7 (0.3–43.1)
Ammonia (mg NH_4/l)[e,b]	log(*x*)	0.05 (0.01–0.6)
Orthophosphate (μg PO_4/l)[e,b]	log(*x*)	60 (10–880)
Total phosphate (μg P/l)[e,b]	log(*x*)	50 (10–490)

[a] Used in physiographic gradient.
[b] Used in global degradation gradient.
[c] Used in hydrological degradation gradient.
[d] Used in morphological gradient.
[e] Used in water quality degradation gradient.

qualitative variables jointly. Consequently, quantitative variables were log-transformed when necessary to better fulfill the PCA normality assumption and the number of meaningful axes was determined by examining the cumulative inertia of the first few axes.

2.2.2. *Metric calculation and standardization*

Ninety-three candidate metrics described in the scientific literature and expected to be impacted by different human-induced degradations were calculated (Table 2). Metrics were based on bio-ecological functional traits (MetFUNC) (fish: 16, diatoms: 17, macrophytes: 14 and macroinvertebrates: 21), on sample taxonomic composition (MetTAX) (fish: 14, diatoms: 14, macrophytes: 3 and macroinvertebrates: 9) or corresponded to previously published indexes (MetIND) (fish: 2, diatoms: 1, macroinvertebrates: 1 and macrophytes: 1). Following Pont et al. (2006, 2007), these metrics were standardized discarding their variability linked to physiography (mainly longitudinal gradient and geology) before examining their responses to human pressures. The first step in the standardization procedure was to model each metric as

$$\text{Biological metric} \sim \text{PG1} + \text{PG2} + \text{PG3} + \text{residuals} \qquad (1)$$

for which the parameters were estimated using CAL-80 minimally disturbed sites. Count data metrics (e.g. number of species) were modelled using log-linear models (McCullagh and Nelder, 1989). Continuous positive and proportional metrics were log-transformed or square-root arcsin transformed, respectively, and modelled using multiple linear models. The predictive reliability of models was assessed using a split-sampling cross-validation method (Harrel, 2001). From these models, predicted metric values were produced for AN-210 sites. The differences between observed and predicted metric values (model residuals) were then standardized (see Pont et al., 2006 for details). In this way, the final values of the metrics were varying according to the intensity of pressures and independently of physiographic gradients. For all the metrics, the independence of the physiographic variables was checked for the minimally disturbed sites of CAL-80 before and after standardization using analysis of variance procedures.

2.2.3. *Biological responses to human pressures*

To consider the combined effect of the different types of pressure, the responses were analysed in three steps for the AN-210 data set. First (step 1), the responses of the metrics to the global pressure gradient were tested. Second (step 2), the responses of these metrics to the presence of an impoundment and to the three gradients of pressure (i.e. water quality, hydrological degradation and morphological degradation) were tested. Finally, the same analysis was carried out for each pressure removing the sites strongly impaired by other types of pressure (step 3). The number of sites impacted by a single pressure in the case of morphological degradations (gp1 = 77; gp4 = 4) and the presence of an impoundment (no = 102; yes = 2) was too low to support statistical tests. Thus, step 3 was not applied for these two pressures.

For each of these steps, in addition to testing the significance of the standardized metric responses to human pressures, the potential to detect human-induced changes was assessed in terms of intensity of the response (i.e. discriminatory efficiency) and sensitivity to changes. For each metric, discriminatory efficiency (DE; Ofenböck et al., 2004) was calculated as the percentage of highly impacted sites (gp4) with metric values below (above) the 5th (95th) percentile of the minimally disturbed sites (gp1) for increasing (decreasing) metrics. The sensitivity of the metric (high,

Table 2
Responses of the 93 metrics to human pressures: discriminatory efficiency (positive or negative) sensitivity. Discriminatory efficiency: the percentage of sites highly impacted (gp4) with metric values less than the extreme percentile (5% for increasing metrics and 95% for decreasing metrics) of the minimally disturbed sites (gp1). (−) for negative and (+) for positive responses. L = low, I = intermediate, H = high sensitivity. BQEs: DI = diatoms, FI = fish, MA = macrophytes, MI = macroinvertebrates; Metric type: MetFUNC = functional trait-based metrics, MetTAX = taxonomy-based metrics, MetIND = Indexes. Units: S = richness, Ni = number of individuals, RF = relative frequency, C = coverage of a taxon/traits, i.e. indices of cumulated density, RA = relative abundance.

Metric	Source	BQE	Type	Description	Global	Water quality		Hydrological		Morphological	Impoundment
					Step 1	Step 2	Step 3	Step 2	Step 3	Step 2	Step 2
M1	(12)	DI	MetFUNC	O_2 intolerant (RA)	57% (−) I	54% (−) I	33% (−) L	19% (−) I		11% (−) H	11% (−)
M2	(12)	DI	MetFUNC	O_2 tolerant (RA)	43% (+) H	53% (+) H	44% (+) H			11% (+) L	11% (+)
M3	(12)	DI	MetFUNC	Aquatic strict (RA)							
M4	(12)	DI	MetFUNC	Terrestrial (RA)				21% (+) L		4% (+) L	
M5	(12)	DI	MetFUNC	Oligosaprobic (RA)							
M6	(12)	DI	MetFUNC	Mesosaprobic (RA)							15% (−)
M7	(12)	DI	MetFUNC	Alphamesosaprobic (RA)		14% (+) L	19% (+) L	11% (+) L			
M8	(12)	DI	MetFUNC	Alphamesotopolysabrobic (RA)	49% (+) H	61% (+) H	48% (+) H			15% (+) L	11% (+)
M9	(12)	DI	MetFUNC	Polysaprobic (RA)							
M10	(12)	DI	MetFUNC	Oligotrophic (RA)							
M11	(12)	DI	MetFUNC	Mesotrophic (RA)	9% (−) H	14% (−) I	7% (−) I				
M12	(12)	DI	MetFUNC	Mesoeutrophic (RA)							33% (−)
M13	(12)	DI	MetFUNC	Trophe indifferent (RA)	43% (−) I	37% (−) I	30% (−) L	28% (−) I		30% (−) L	22% (−)
M14	(12)	DI	MetFUNC	Sensitive N-autotrophic (RA)		10% (−) L					
M15	(12)	DI	MetFUNC	Tolerant N-autotrophic (RA)							
M16	(12)	DI	MetFUNC	Facultative N-heterotrophic (RA)	43% (+) H	37% (+) H	26% (+) H				
M17	(12)	DI	MetFUNC	Obligatory N-heterotrophic (RA)	20% (+) L			18% (+) L		15% (+) L	15% (+)
M18		DI	MetTAX	S	34% (+) H	20% (+) I	22% (+) L				
M19	(7)	DI	MetTAX	Evenness ($E = e^H/S$)							
M20		DI	MetTAX	Ni							
M21	(10)	DI	MetTAX	Shannon diversity (H)	9% (+) L	10% (+) I					
M22	(4)	DI	MetIND	IBD (French Diatom Index)	80% (−) H	61% (−) I	52% (−) I	21% (−) I	7% (−) L	7% (−) H	19% (−)
M23	(5)	FI	MetFUNC	Insectivorous (S)							63% (−)
M24	(5)	FI	MetFUNC	Insectivorous (RA)	37% (−) L			35% (−) L		30% (−) L	48% (−)
M25	(5)	FI	MetFUNC	Omnivorous (S)							
M26	(5)	FI	MetFUNC	Omnivorous (RA)	57% (+) L	25% (+) L		14% (+) I		19% (+) H	11% (+)
M27	(5)	FI	MetFUNC	O_2 intolerant (S)	71% (−) L			46% (−) I	18% (−) L	48% (−) I	85% (−)
M28	(5)	FI	MetFUNC	O_2 intolerant (RA)	91% (−) H	61% (−) I	37% (−) I	32% (−) I		37% (−) H	41% (−)
M29	(5)	FI	MetFUNC	O_2 tolerant (S)	74% (+) H	49% (+) L	52% (+) L			26% (+) L	33% (+)
M30	(5)	FI	MetFUNC	O_2 tolerant (RA)	89% (+) H	71% (+) I	52% (+) I	32% (+) L		48% (+) H	44% (+)
M31	(5)	FI	MetFUNC	Habitat intolerant (S)	66% (−) L	41% (−) L		35% (−) L		41% (−) L	52% (−)
M32	(5)	FI	MetFUNC	Habitat intolerant (RA)	14% (−) L	3% (−) L					
M33	(5)	FI	MetFUNC	Habitat tolerant (S)	69% (+) H	53% (+) H	52% (+) I				33% (+)
M34	(5)	FI	MetFUNC	Habitat tolerant (RA)	91% (+) H	56% (+) I	41% (+) I	39% (+) I		44% (+) H	41% (+)
M35	(5)	FI	MetFUNC	Rheopar (S)	60% (−) L			28% (−) L		44% (−) L	70% (−)
M36	(5)	FI	MetFUNC	Rheopar (RA)	91% (−) H	56% (−) I	52% (−) I	39% (−) L		52% (−) H	63% (−)
M37	(5)	FI	MetFUNC	Lithophilic (S)	63% (−) L			32% (−) L		48% (−) L	74% (−)
M38	(5)	FI	MetFUNC	Lithophilic (RA)	91% (−) H	54% (−) H	52% (−) H	33% (−) L		41% (−) L	30% (−)
M39		FI	MetTAX	S							
M40	(7)	FI	MetTAX	Evenness ($E = e^H/S$)							
M41		FI	MetTAX	Ni							
M42	(10)	FI	MetTAX	Shannon diversity (H)							15% (−)
M43	(1)	FI	MetIND	IPR (French Fish Index)	49% (−) L			32% (−) I	21% (−) I	22% (−) L	22% (−)
M44	(9)	FI	MetIND	EFI (European Fish Index)	74% (−) L	47% (−) L		21% (−) I		15% (−) L	15% (−)
M45	(3)	MA	MetFUNC	Water quality intolerant (S)	9% (−) I			2% (−) I		7% (−) L	
M46	(3)	MA	MetFUNC	Water quality intolerant (C)				2% (−) H			
M47	(3)	MA	MetFUNC	Water quality tolerant (S)		41% (+) I	56% (+) I				19% (−)
M48	(3)	MA	MetFUNC	Water quality tolerant (C)		31% (+) L	26% (+) L				
M49	(3)	MA	MetFUNC	Amphibious (S)			37% (+) I	25% (−) I			22% (−)

M50	(3)	MA	MetFUNC	Amphibious (C)				30% (−) I	29% (−) L		37% (−)
M51	(3)	MA	MetFUNC	Aquatic strict (S)			15% (+) L				33% (−)
M52	(3)	MA	MetFUNC	Aquatic strict (C)		39% (+) I	56% (+) I				
M53	(3)	MA	MetFUNC	Helophytes (S)	43% (+) L	41% (+) I					
M54	(3)	MA	MetFUNC	Helophytes (C)							
M55	(3)	MA	MetFUNC	Intolerant species (S)							26% (+)
M56	(3)	MA	MetFUNC	Intolerant species (C)	6% (+) L						7% (+)
M57	(3)	MA	MetFUNC	Tolerant species (S)		37% (+) H	52% (+) H				26% (−)
M58	(3)	MA	MetFUNC	Tolerant species (C)			30% (+) L				
M59		MA	MetTAX	S		27% (+) H	33% (+) H				22% (−)
M60	(7)	MA	MetTAX	Evenness ($E = e^H/S$)							
M61	(10)	MA	MetTAX	Shannon diversity (H)			41% (+) L				22% (−)
M62	(6)	MA	MetIND	IBMR (French Macrophyte Index)	20% (−) H	15% (−) I	15% (−) I			11% (−) L	7% (−)
M63	(11)	MI	MetFUNC	Deposit feeder (RF)			41% (−) I				
M64	(11)	MI	MetFUNC	Shredder (RF)							
M65	(11)	MI	MetFUNC	Maximum size < 2.5 cm (RF)					11% (−) L		
M66	(11)	MI	MetFUNC	Maximum size > 8 cm (RF)							
M67	(11)	MI	MetFUNC	Life cycle < 1 year (RF)	49% (−) L	47% (−) I	33% (−) I			19% (−) L	11% (−)
M68	(11)	MI	MetFUNC	Lyfe cycle > 1 year (RF)	29% (+) L	36% (+) I	33% (+) L				
M69	(11)	MI	MetFUNC	Number of cycles per year < 1 (RF)	40% (−) H	25% (−) I	22% (−) I			30% (−) L	19% (−)
M70	(11)	MI	MetFUNC	Number of cycles per year > 1 (RF)	29% (+) L	22% (+) L				30% (+) L	19% (+)
M71	(11)	MI	MetFUNC	Ovoviviparity (RF)	86% (+) H	68% (+) I	52% (+) I	23% (+) I		37% (+) L	26% (+)
M72	(11)	MI	MetFUNC	Aquatic passive dispersal (RF)	57% (+) H	51% (+) I	48% (+) L			22% (+) L	
M73	(11)	MI	MetFUNC	Aerial active dispersal (RF)	77% (−) H	75% (−) I	81% (−) I	21% (−) L		48% (−) L	30% (−)
M74	(11)	MI	MetFUNC	Gravel (RF)	51% (−) L					48% (−) L	59% (−)
M75	(11)	MI	MetFUNC	Sand (RF)	46% (−) I	36% (−) L	56% (−) L	19% (−) L		33% (−) L	30% (−)
M76	(11)	MI	MetFUNC	Microphytes (RF)	40% (+) H	29% (+) I	30% (+) I			11% (+) L	
M77	(11)	MI	MetFUNC	Mud (RF)							
M78	(11)	MI	MetFUNC	Xenosaprobic (RF)	26% (−) H	25% (−) L				19% (−) L	11% (−)
M79	(11)	MI	MetFUNC	Oligosaprobic (RF)	49% (−) H	37% (−) L		12% (−) I		22% (−) L	19% (−)
M80	(11)	MI	MetFUNC	Polysaprobic (RF)	26% (+) L	17% (+) L				26% (+) L	26% (+)
M81	(11)	MI	MetFUNC	Oligotrophic (RF)	54% (−) H	39% (−) I		11% (−) I		26% (−) L	15% (−)
M82	(11)	MI	MetFUNC	Mesotrophic (RF)							19% (−)
M83	(11)	MI	MetFUNC	Eutrophic (RF)	43% (+) H	29% (+) L		12% (+) L		22% (+) L	15% (+)
M84		MI	MetTAX	S							
M85	(7)	MI	MetTAX	Evenness ($E = e^H/S$)							
M86		MI	MetTAX	Ni							7% (−)
M87	(10)	MI	MetTAX	Shannon diversity (H)	34% (−) L						15% (−)
M88		MI	MetTAX	Ephemeroptera, Plecoptera, Trichoptera taxa (%)	23% (−) L	12% (−) L				7% (−) L	7% (−)
M89		MI	MetTAX	Plecoptera taxa (%)	20% (−) H	19% (−) I	11% (−) I	16% (−) H		22% (−) L	
M90	(8)	MI	MetTAX	GoldInv							33% (+)
M91	(13)	MI	MetIND	IBGN (French Macroinvertebrate Index)	60% (−) I	41% (−) L	33% (−) L	16% (−) L		37% (−) L	30% (−)
M92	(2)	MI	MetIND	BMWP (Biological Monitoring Working Party)	26% (−) L	32% (−) L		16% (−) I		37% (−) H	22% (−)
M93	(2)	MI	MetIND	ASPT (Average Score per Taxa)	80% (−) H	71% (−) I	44% (−) L	37% (−) I		52% (−) H	22% (−)

Sources: (1) AFNOR (2004b), (2) Armitage et al. (1983), (3) Christian Chauvin (pers. comm.), (4) Coste et al. (2009), (5) Haury et al. (2006), (6) Pont et al. (2006), (7) Pielou (1966), (8) Pinto et al. (2004), (9) Pont et al. (2007), (10) Shannon and Weaver (1949), (11) Tachet et al. (2006), (12) van Dam et al. (1994), (13) Genin et al. (2003).

intermediate or low, respectively) was defined by the lowest level of pressure (gp2, gp3 or gp4, respectively) for which metric values were significantly different from metric values for minimally disturbed conditions (gp1), i.e. the lowest level of pressure detected by the metric. To compare BQEs, weighted means of sensitivity (WS) were calculated for each BQE, allocating weight for the three situations: respectively 1, 2 and 3 for low, intermediate or high sensitivity. A WS close to 3 indicates high sensitivity of the metric to the pressure, i.e. response to low level of pressure, while a WS close to 1 indicates low sensitivity, i.e. response only to high level of pressure. Kruskal–Wallis non-parametric post hoc tests were used to determine the effect of single and combined human pressures on BQE metrics. Statistical analyses were implemented using R software (version 2.10.1).

3. Results

3.1. Physiographic gradients and anthropogenic pressure indices

The first three axes of the physiographic variable analysis accounted for 53.5% of the total inertia with PG1, PG2 and PG3 explaining 31.6, 12.5 and 9.3%, respectively. PG1 was related to a longitudinal gradient, which increased with altitude and mean slope and decreased with mean width and catchment area. PG2 was related to the same variables but did not suggest a clear interpretation. PG3 was related to geological types from siliceous to calcareous.

The global synthetic pressure index explained 20.1% of the total inertia, decreasing when oxygen saturation increased and increasing with all the other chemical (e.g. total phosphorus, nitrate) and hydro-morphological degradation variables (from “no” to “high”; Table 1). The synthetic indexes corresponding to water quality, hydrological and morphological degradations (the first axes of the analyses including only water quality, hydrological and morphological variables) accounted for 55.9%, 23.6% and 22.5%, respectively, of the total inertia and were related to an increase in disturbances of the associated variables.

3.2. Physiography and standardized biological metrics

For the CAL-80 undisturbed sites, before standardization, sixty out the 93 metrics tested varied significantly (ANOVA p-value > 0.05) along the longitudinal gradient (PG1). Not surprisingly, the 33 unvarying metrics included the four Ecological quality ratio (EQR) tested (IBD: M22, IPR: M43, EFI: M44, IBGN: M91). As expected, ANOVA procedures did not reveal residual physiographic effect for the 93 standardized metrics (e.g. percentage of the Plecoptera taxa: M89; Fig. 2). In addition, the cross-validation showed that the models were stable. The root mean squared error (RMSE) and Spearman rank correlation between predictions and observations were coherent between CAL-80 and the 200 resamples, and deviations rarely exceeded 15% of the initial statistics.

3.3. Biological responses to anthropogenic pressure gradients

Four types of metric were identified regarding its responses to pressures (examples in Fig. 3 and exhaustive results in Table 2): metrics not impacted by pressures (i.e. no response, 27 metrics), decreasing metrics (i.e. negative response, 34 metrics), increasing metrics (i.e. positive response, 26 metrics), and decreasing/increasing metrics depending on the type of pressure (six macrophyte metrics).

3.3.1. Global degradation

From the 51 metrics responding significantly to the global pressure gradient (significance rates fish: 73, macroinvertebrates: 68,

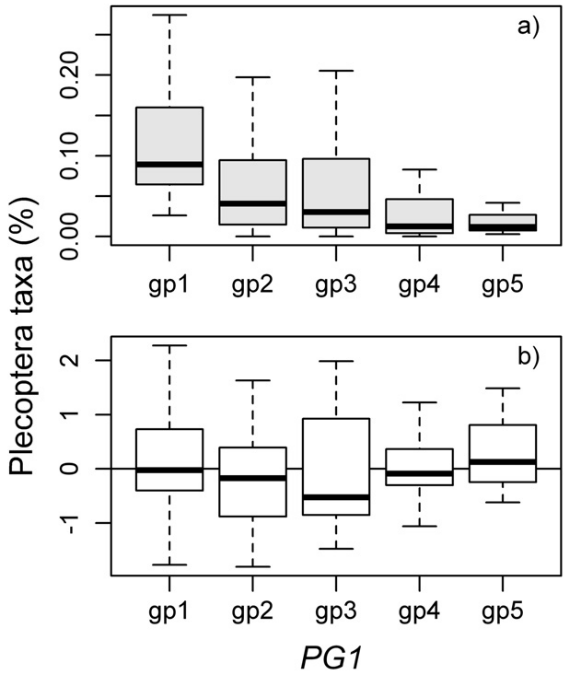


Fig. 2. Boxplots of the percentage of Plecoptera taxa along the first physiographic gradient (PG1) for undisturbed sites (CAL-80) before (a) and after (b) standardization. For illustrative purposes the longitudinal gradient PG1 was represented by five classes (from small mountain streams (gp1) to large lowland rivers (gp5)).

diatoms: 45, macrophytes: 22% of the total number of metrics for each BQE), 46 had a discriminatory efficiency over 20% (Table 2). The strongest (DE = 91%) and most sensitive (High) responses were observed for the four fish MetFUNCs: M29, M34, M36, M38. Most of the 16 fish metrics showed a strong response (DE median value 70%; range: 14–91%), half of them to a low pressure level (High: 44%, Low: 56%; WS = 1.9). The ten diatom metrics and the 21 macroinvertebrate metrics showed, on average, similar response intensities (DE medians: 43% and ranges diatoms: 9–80%, macroinvertebrates: 20–86%) and mostly high sensitivity for diatoms (High: 60%, Intermediate: 20%, Low: 20%; WS = 2.4) and macroinvertebrates (High: 52%, Intermediate: 10%, Low: 38%; WS = 2.1). The four macrophyte metrics presented rather weak responses (DE range, 6–43%) and low sensitivity (High: one, Intermediate: one, Low: two metrics; WS = 1.8).

3.3.2. Water quality degradation

Forty-eight metrics responded significantly to water quality degradation gradients (significance rates macroinvertebrates: 61, diatoms: 50, fish: 50, macrophytes: 39%) and 39 had a DE over 20% (Table 2). The strongest (DE = 61, 75, 71 and 71%, respectively) and most sensitive responses were observed for three MetFUNCs (M8 (High), M30 (Intermediate) and M73 (Intermediate)) and the macroinvertebrate index ASPT (M93 (Intermediate)). Most of the fish metrics presented rather low sensitivity (High: 18%, Intermediate: 36%, Low: 45%; WS = 1.7), but the strongest responses (DE median 53%, range: 3–71%). Whereas macroinvertebrate metrics showed intermediate or low sensitivity (Intermediate: 53%, Low: 47%; WS = 1.5), most of the diatom (High: 27%, Intermediate: 55%, Low: 18%; WS = 2.1) and macrophyte (High: 29%, Intermediate: 57%, Low: 14%; WS = 2.2) metrics were highly sensitive. However, their responses presented similar median intensities with wider ranges for diatoms and macroinvertebrates than for macrophytes (DE medians (ranges) macrophytes: 37 (15–41%), diatoms: 37 (10–61%) and macroinvertebrates: 36% (12–75%)).

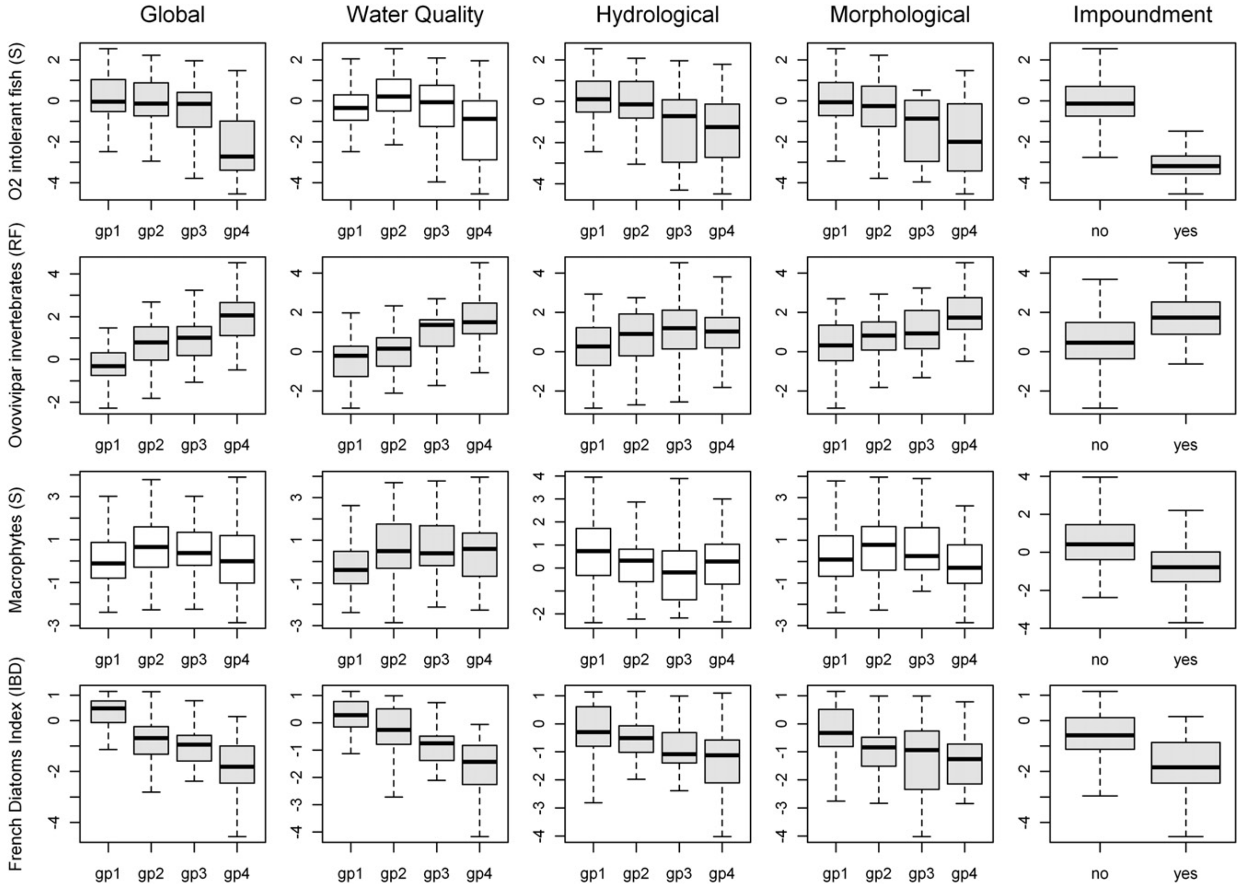


Fig. 3. Boxplots of the responses of four metrics to human-induced perturbations, i.e. global, water quality, hydrological, morphological degradation gradients and the presence of an impoundment. White boxes (respectively, grey boxes) for nonsignificant (significant) responses. Pressure gradients represented by four classes from minimally disturbed sites (gp1) to highly impacted sites (gp4).

Thirty-eight metrics still responded to water quality deterioration when considered as a single pressure (significance rates macrophytes: 56, diatoms: 41, macroinvertebrates: 39, fish: 32%). DE median values (macrophytes: 30, diatoms: 35, macroinvertebrates: 37, fish: 52%) and WS (macrophytes: 2, diatoms: 2, fish: 2, macroinvertebrates: 1.6) remained nearly unchanged for the four BQEs. The strongest (DE = 52, 52 and 81%, respectively) and most sensitive responses were observed for three MetFUNCs (M38 (High), M57 (High) and M73 (Intermediate)).

3.3.3. Hydrological degradation

Thirty-three metrics responded to the hydrological degradation gradient (significance rates, fish: 59, macroinvertebrates: 32, diatoms: 27, macrophytes: 22%) and 20 had a DE over 20% (Table 2). The strongest (DE = 16 and 81%, respectively) and most sensitive responses were observed for M89 (High) and M28 (Intermediate). Metric responses were generally weak (BQEs DE medians < 32% and BQE DE range: 2–46%) with low sensitivity (fish WS = 1.5, macroinvertebrates WS = 1.7, diatoms WS = 1.5) except for the macrophytes (WS = 2.3; only two highly sensitive metrics: M46 (macrophytes) and M89 (macroinvertebrates); Table 2). Only five metrics still showed significant responses when considering hydrological degradation as a single pressure (M50 (macrophytes), M22 (diatoms), M65 (macroinvertebrates), M27 and M43 (fish); Table 2) and with rather low sensitivity (fish WS = 1.5, macroinvertebrates WS = 1, diatoms WS = 1, macrophytes WS = 1).

3.3.4. Morphological degradation

Forty-two metrics responded to the morphological degradation gradient (significance rates, fish: 64, macroinvertebrates: 61, diatoms: 32, macrophytes: 11%) and 28 had a DE over 20% (Table 2). The strongest (DE = 52%) and most sensitive (High) responses were observed for M36 and M93. Most of the fish metrics showed intermediate intensity responses (DE median 41% and range: 15–52%) and one-third were highly sensitive (High: 36%, Intermediate: 7%, Low: 57%; WS = 1.8). Most of the macroinvertebrate metrics showed weak responses (DE median 26%, range: 7–52%) to a high level of morphological degradation (High: 11%, Intermediate: 89%; WS = 1.2). Diatom and macrophyte metrics mainly showed very weak responses (DE medians < 11% and range 4–30%). While all the macrophyte metrics showed low sensitivity (WS = 1), one-third of diatom metrics were highly sensitive (High: 29%, Low: 71%; WS = 1.6).

3.3.5. Impoundment

Fifty-five metrics detected the presence of an impoundment (significance rates, fish: 77, macroinvertebrates: 65, macrophytes: 56, diatoms: 36%) and 32 had a DE over 20% (Table 2). The strongest responses were observed for the two fish MetFUNCs M28 and M37 (DE = 85 and 74%, respectively). Most of the fish metrics responded with medium intensities (DE median 41%, range: 11–85%), while most of the macroinvertebrate, macrophyte and diatom metrics responded weakly (DE median (range) 19% (7–59%), 22% (7–37%), and 15% (11–33%), respectively).

3.3.6. Comparison of metric types

All the MetIND, two-thirds of the MetFUNC and one-third of the MetTAX responded to the global degradation gradient. MetIND and MetFUNC showed the strongest responses with median discriminatory efficiency (DE) respectively equal to 60% (range: 20–80%) and 49% (range: 6–91%). Half of the MetFUNCs were highly sensitive (High: 52%, Intermediate: 10%, Low: 38%, WS = 2.1) while the MetINDs were globally less sensitive (High: 43%, Intermediate: 14%, Low: 43%; WS = 2). Most of the MetTAXs showed weak responses and low sensitivity (DE 23% (9–34%); High: 40%, Low: 60%; WS = 1.8). Similar results were observed for the specific pressure gradients, i.e. MetIND and MetFUNC responded more frequently and more strongly than MetTAX.

4. Discussion

The main purpose of this study was to test whether river assemblage responses to human-induced changes were similar among macrophytes, diatoms, fish and macroinvertebrates. More particularly, the potential to detect human-induced changes was compared in terms of intensity of the response (i.e. discriminatory efficiency) and sensitivity to changes (i.e. first significant responses occurring along pressure gradients). Metrics were transformed beforehand to retain only the proportion of the signal related to human-induced changes. Sixty-six out of the 93 transformed metrics detected at least one of the five anthropogenic pressures considered. The strongest efficiency and sensitivity were observed for MetFUNC and MetIND. Also, BQEs responded differently, depending on the type of human pressure. As pointed out by Johnson et al. (2006b), discriminatory efficiency and sensitivity varied noticeably among individual metrics for a given BQE.

4.1. Taking into account physiography in the analysis

As advised by Pont et al. (2007), the calibration data set covers mainly the range of physiographic diversity of the French territory (large rivers are not considered in this study). Before transformation, two-thirds of the tested metrics varied significantly with the physiographic variables when only considering weakly disturbed sites. These results are in agreement with previous studies showing that local and regional physiographic factors are major drivers of change in BQE structure and function (Hughes et al., 2009; Johnson et al., 2006b; Logez et al., 2010). Hence, this source of variability should be taken into account before considering biological responses strictly due to human-induced stressors.

Moreover, in this study, biological assemblages showed different responses to the physiographic gradients. Indeed, BQEs are known to be related differently to their environment and thus common important physiographic parameters and pressure descriptors were selected in this study. Although introducing uncertainty into the models, this compromise was indispensable for comparison.

All the standardized metrics tested were independent of physiography for undisturbed sites. These results indicate that the method used is appropriate to discard the portion of metric variability related to the direct effect of natural phenomena while still considering the river as a continuum (Vannote et al., 1980), i.e. not splitting the data set into different river types.

4.2. Intensity and sensitivity of BQE responses to general degradation

As in previous studies, metric responses were stronger overall for global degradation than for specific pressures (Hering et al., 2006). In addition, the present study demonstrates that metrics are more sensitive to global pressure. Nevertheless, sensitivity and intensity of metric responses to human pressures fluctuated among biological groups. Diatom and macroinvertebrate metrics appear to be more sensitive to the degradation of the overall condition of the river than fish metrics and reacted to lower levels of pressure. However, fish metrics presented the strongest intensities. These differences may be partly related to the migratory capacities of fish and their longer life cycles. Consequently, as long as favourable habitats and conditions are accessible for fish, changes will remain undetected by metrics. When they are no longer accessible, fish metrics will show dramatic responses resulting in strong responses to strong perturbations. Conversely, short-life-cycle and sedentary organisms such as benthic diatoms will be impacted by a lower level of pressure as soon as local favourable conditions are degraded. The less sensitive but more intense responses of fish metrics would be better adapted to detecting high modifications or the first results of restoration measures while diatom and macroinvertebrate metrics would be more useful in detecting the first impacts of degradation and more advanced stages of restoration.

Compared to the other BQEs, fewer macrophyte metrics responded significantly and responses were weaker and less sensitive. These differences may be due to the positive/negative responses of the macrophyte metrics since all macrophyte metrics showing contrasted responses to specific pressures are insensitive to global degradation. For example, M47, i.e. the number of water quality-tolerant macrophyte species, were positively impacted by a diminution of the water quality (DE = 41%, intermediate response) but negatively impacted by the presence of an impoundment (DE = 19%). Therefore, in multi-impacted sites, losses caused by a pressure could be compensated by benefits related to other pressures revealing possible antagonistic effects. Such metrics could be particularly useful in detecting impacts related to different types of pressure but could be confounding when assessing general degradation of multi-impacted sites.

4.3. BQE responses to hydrological, morphological and water quality degradations

Among specific pressures, water quality degradations resulted in the strongest responses in term of intensity and sensitivity. In agreement with previous studies, the four BQEs showed significant responses to water quality degradation (Hering et al., 2006; Johnson et al., 2006b; Johnson and Hering, 2009; Justus et al., 2010). In this study, diatom and macrophyte metrics were more sensitive to water quality (response to low to moderate levels of water quality degradation) than fish and macroinvertebrate metrics (response to moderate to high levels of water quality degradation). However, as in Johnson and Hering (2009), the strongest responses were observed for fish metrics.

Contrary to Hughes et al. (2009), who showed that macroinvertebrates and macrophytes were more impacted by water velocity changes and fish by physical disturbance, the present results suggest that fishes are the most impacted by hydrological perturbations followed by macrophytes and diatoms. Nonetheless, responses were rather weak (DE median <32%) and significant for a high pressure level for all groups (WS < 1.7) except macrophytes (WS = 2.3). Besides, fish and macroinvertebrate metrics showed the strongest responses to morphological degradations. Although fish, diatom and macroinvertebrate metrics seem to be the most sensitive metrics to this pressure, responses generally occurred for high degradation. In previous studies, hydrological and morphological perturbations were generally combined into a single pressure gradient identified as habitat degradation. The present results tend to be very similar to those of Hering et al. (2006) showing that macroinvertebrates and fish responded to reach scale hydro-morphological gradients and contrast with those of Johnson et al. (2006b) showing that fish and macrophyte metrics showed a more substantial response to general habitat alteration than either diatom or macroinvertebrate metrics. These differences may be explained by the authors' choice of variables describing pressures and the analysis settings.

The four BQE metrics were affected by the presence of an impoundment and the highest responses were observed for fish metrics. The results of this study confirmed the particular ecological impact of impoundments with relatively strong responses of the four BQEs to this pressure. Indeed, this type of river modification is known to strongly alter both water quality and hydro-morphological conditions upstream and downstream of a weir or a dam (Baxter, 1977; Feld et al., 2011).

4.4. Detecting combined and single pressure effects

For both water quality and hydrological degradations, the number of significant responses decreased sharply from pressure types (step 2) to single pressure analysis (step 3). For instance, for the fish metrics, the significance rate fell from 59 to 9% for hydrological pressure and from 50 to 32% for water quality degradation. In addition, the same pattern as in step 2 was observed in step 3, i.e. better responses in terms of intensity and sensitivity to water quality than to hydrological perturbations.

The intensity and sensitivity of the four BQE responses to water quality degradation were nearly unchanged when removing sites strongly impacted by hydrological or morphological degradations (from step 2 to step 3). This result suggests that the effects on BQE metrics observed in step 2 were mainly due to water quality degradation and not to a combined effect with hydro-morphological degradation.

When sites strongly impacted by water quality or morphological degradations were not considered (from step 2 to step 3), the responses of fish metrics to hydrological pressure gradient were nearly unchanged, whereas macrophyte, macroinvertebrate and diatom metrics were less sensitive and showed weaker responses to this gradient. Therefore, the effects observed on fish metrics in step 2 appear to be mainly due to hydrological changes and not to a combined effect with water quality and/or morphological degradations. On the other hand, macrophyte, macroinvertebrate and diatom metric responses probably largely resulted from the impact of associated water quality and/or morphological degradations. These results raise new issues about the relation between pressures. Unfortunately, the understanding of the combined or cumulated effect of several types of pressure on river aquatic assemblages is typically poor (Pont et al., 2007) and questions such as the existence of cumulative or multiplicative (i.e. interaction) effects remain unanswered.

4.5. Choosing metric types

Indexes and functional trait-based metrics were generally more sensitive and showed stronger responses to pressures than taxonomy-based metrics. Indeed, several authors have demonstrated that ecological and biological functional traits are well-adapted for large-scale approaches (Statzner et al., 2001) and are able to integrate more general phenomena than taxonomy-based metrics (Dolédec et al., 2006). Also, the five indexes tested in this study (Table 2: M22, M43, M44, M62, M93) showed strong responses to global degradation. Apart from M43 (IPR: French Fish Index) and M62 (IBMR: French Macrophyte Index), which did not detect water quality and morphological perturbations, respectively, indexes were significantly affected by the three specific pressure gradients and the presence of an impoundment. These results clearly support the use of ecological and biological functional trait metrics to build multi-metric indexes to assess river biotic integrity. We advocate that the selection of the metric to monitor the effects of stressors of interest should not only focus on the BQE, but also on the nature of the metric (i.e. underlying processes, types, and units).

5. Conclusions

This study demonstrates that the two main sources of variability in biological assemblages (physiography diversity and anthropogenic disturbances) should be distinguished a priori when looking at the impacts of human-induced stressors. Also, when selecting the best BQEs or metrics to detect stressor impacts, particular care should be taken when selecting the type of metric to study. Indeed, indexes and functional trait-based metrics tend to detect human-induced changes better (stronger responses to a lower level of pressure) than taxonomy-based metrics. However, to expand this analysis, knowledge of biological and ecological traits needs to be improved, in particular for macrophytes.

This study shows that the four BQE metrics are impacted differently by pressures, even if the responses vary from one metric to another within a given group. Metric selection within a group might have as much importance in precisely detecting the impact of human-induced change as selecting a BQE. More generally, global and water quality degradations of the river appear to be better detected by BQE metrics than (in decreasing order) impoundments, morphological degradations and hydrological degradations. Finally, given the present results, hydrological degradation effects will likely be confounded with water quality and morphological degradation effects on the biota if multi-impacted sites are not removed from the analysis.

As river assessment research is turning towards multi-metric tools, it is of prime importance to be able to answer the following question before including metrics in indexes: Do the different types of pressure have additive, multiplicative or opposite effects? Furthermore, this study has analysed the influence of the physiography on metrics for undisturbed conditions, but we believe that complex interactions exist between human pressure effects and the environmental diversity, i.e. the responses of aquatic assemblages to human pressure will be different depending on the physiography. Such interaction effects on BQE responses have not been sufficiently investigated to date and are needed to assess the ecological impacts towards restoration of water bodies.

Acknowledgements

This paper is part of the WISER project (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract no. 226273), www.wiser.eu. It has been supported by the collaborative project "HYNES" between Irstea and the Research and Development Department of the French Electric Company (EDF). We thank Philippe Usseglio-Polatera, Martial Ferreol and the LHQ-ONEMA/Irstea research team for their support and contributions to the macroinvertebrate and diatom data.

References

AFNOR, 2003. Qualité de l'eau – Détermination de l'indice biologique macrophytique en rivière (IBMR). Association française de normalisation, Norme homologuée T90-395.
AFNOR, 2004a. Qualité de l'eau – Détermination de l'indice poissons rivières (IPR). Association française de normalisation, Norme homologuée T90-344.
AFNOR, 2004b. Qualité écologique des milieux aquatiques. Qualité de l'eau – Détermination de l'indice biologique global normalisé (IBGN). Association française de normalisation, Norme homologuée T90-350.
AFNOR, 2007. Qualité de l'eau – Détermination de l'Indice Biologique Diatomées (IBD). Association française de normalisation, Norme homologuée NF T90-354.
Archaimbault, V., Usseglio-Polatera, P., Garric, J., Wasson, J.G., Babut, M., 2010. Assessing pollution of toxic sediment in streams using bio-ecological traits of benthic macroinvertebrates. Freshw. Biol. 55, 1430–1446.

Armitage, P.D., Moss, D., Wright, J.F., Furse, M.T., 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide-range of unpolluted running-water sites. Water Res. 17, 333–347.
Baxter, R.M., 1977. Environmental effects of dams and impoundments. Annu. Rev. Ecol. Syst. 8, 255–283.
Besse-Lototskaya, A., Verdonschot, P.F.M., Coste, M., Van de Vijver, B., 2011. Evaluation of European diatom trophic indices. Ecol. Indic. 11, 456–467.
Carpenter, K.D., Waite, I.R., 2000. Relations of habitat-specific algal assemblages to land use and water chemistry in the Willamette Basin. Oregon. Environ. Monit. Assess. 64, 247–257.
Coste, M., Boutry, S., Tison-Rosebery, J., Delmas, F., 2009. Improvements of the Biological Diatom Index (BDI): description and efficiency of the new version (BDI-2006). Ecol. Indic. 9, 621–650.
D'Ambrosio, J.L., Williams, L.R., Witter, J.D., Ward, A., 2009. Effects of geomorphology, habitat, and spatial location on fish assemblages in a watershed in Ohio, USA. Environ. Monit. Assess. 148, 325–341.
Dolédec, S., Phillips, N., Scarsbrook, M., Riley, R.H., Townsend, C.R., 2006. Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. J. N. Am. Benthol. Soc. 25, 44–60.
EC, 2000. 2000/60/EC, Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off. J. Eur. Commun. 327, 1–72.
Fausch, K.D., Lyons, J., Karr, J.R., Angermeier, P.L., 1990. Fish communities as indicators of environmental degradation. In: Biological Indicators of Stress in Fish. American Fisheries Society Symposium, vol. 8, Bethesda, Maryland, USA, pp. 123–144.
Feld, C.K., Birk, S., Bradley, D.C., Hering, D., Kail, J., Marzin, A., Melcher, A., Nemitz, D., Pedersen, M.L., Pletterbauer, F., Pont, D., Verdonschot, P.F.M., Friberg, N., 2011. From natural to degraded rivers and back again: a test of restoration ecology theory and practice. Adv. Ecol. Res. 44, 119–209.
Fore, L.S., Grafe, C., 2002. Using diatoms to assess the biological condition of large rivers in Idaho (USA). Freshw. Biol. 47, 2015–2037.
French Water Agency, 2000. Système d'évaluation de la qualité des cours d'eau. Rapport de présentation SEQ-Eau. Etudes des Agences de l'Eau 64, 59 pp.
Genin, B., Chauvin, C., Ménard, F., 2003. Cours d'eau et indices biologiques pollution, méthodes, IBGN, 2nd edition. Educagri, Dijon.
Harrel, F.E., 2001. Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer, New York.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS136: a *k*-means clustering algorithm. Appl. Stat. 28, 100–108.
Haury, J., Peltre, M.C., Trémolières, M., Barbe, J., Thiebaut, G., Bernez, I., Daniel, H., Chatenet, P., Haan-Archipof, G., Muller, S., Dutartre, A., Laplace-Treyture, C., Cazaubon, A., Lambert-Servien, E., 2006. 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 570, 153–158.
Heino, J., Paavola, R., Virtanen, R., Muotka, T., 2005. Searching for biodiversity indicators in running waters: do bryophytes, macroinvertebrates, and fish show congruent diversity patterns? Biodivers. Conserv. 14, 415–428.
Heino, J., 2010. Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? Ecol. Indic. 10, 112–117.
Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K., Verdonschot, P.F.M., 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshw. Biol. 51, 1757–1785.
Hill, M.O., Smith, A.J.E., 1976. Principal component analysis of taxonomic data with multi-state discrete characters. Taxon 25, 249–255.
Hughes, S.J., Santos, J.M., Ferreira, M.T., Caraça, R., Mendes, A.M., 2009. Ecological assessment of an intermittent Mediterranean river using community structure and function: evaluating the role of different organism groups. Freshw. Biol. 54, 2383–2400.
Johnson, R.K., Hering, D., Furse, M.T., Verdonschot, P.F.M., 2006a. Indicators of ecological change: comparison of the early response of four organism groups to stress gradients. Hydrobiologia 566, 139–152.
Johnson, R.K., Hering, D., Furse, M.T., Clarke, R.T., 2006b. Detection of ecological change using multiple organism groups: metrics and uncertainty. Hydrobiologia 566, 115–137.
Johnson, R.K., Hering, D., 2009. Response of taxonomic groups in streams to gradients in resource and habitat characteristics. J. Appl. Ecol. 46, 175–186.
Justus, B.G., Petersen, J.C., Femmer, S.R., Davis, J.V., Wallace, J.E., 2010. A comparison of algal, macroinvertebrate, and fish assemblage indices for assessing low-level nutrient enrichment in wadeable Ozark streams. Ecol. Indic. 10, 627–638.
Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.
Lacoul, P., Freedman, B., 2006. Environmental influences on aquatic plants in freshwater ecosystems. Environ. Rev. 14, 89–136.
Logez, M., Pont, D., Ferreira, M.T., 2010. Do Iberian and European fish faunas exhibit convergent functional structure along environmental gradients? J. N. Am. Benthol. Soc. 29, 1310–1323.
Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., Randers, J., 2005. The Living Planet Index: using species population time series to track trends in biodiversity. Philos. Trans. R. Soc. B: Biol. Sci. 360, 289–295.
Lorenz, A., Hering, D., Feld, C.K., Rolauffs, P., 2004. A new method for assessing the impact of hydromorphological degradation on the macroinvertebrate fauna of five German stream types. Hydrobiologia 516, 107–127.
McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, 2nd edition. Chapman and Hall, London.
Oberdorff, T., Pont, D., Hugueny, B., Chessel, D., 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. Freshw. Biol. 46, 399–415.
Ofenböck, T., Moog, O., Gerritsen, J., Barbour, M., 2004. A stressor specific multimetric approach for monitoring running waters in Austria using benthic macroinvertebrates. Hydrobiologia 516, 251–268.
Pantle, R., Buck, H., 1955. Die biologische Überwachung der Gewässer und die. Darstellung der Ergebnisse. Gas Wasserfach 96, 604.
Petts, G.E., 1989. Historical analysis of fluvial hydrosystems. In: Petts, G.E., Muller, H., Roux, A.L. (Eds.), Historical Change of Large Alluvial Rivers: Western Europe. Wiley, Chichester, pp. 1–18.
Pielou, E.C., 1966. The measurement of diversity in different types of biological collections. J. Theor. Biol. 13, 131–144.
Pinto, P., Rosado, J., Morais, M., Antunes, I., 2004. Assessment methodology for southern siliceous basins in Portugal. Hydrobiologia 516, 191–214.
Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., Schmutz, S., 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. J. Appl. Ecol. 43, 70–80.
Pont, D., Hugueny, B., Rogers, C., 2007. Development of a fish-based index for the assessment of "river health" in Europe: the European Fish Index (EFI). Fish. Manag. Ecol. 14, 427–439.
Riis, T., Sand-Jensen, K., Vestergaard, O., 2000. Plant communities in lowland Danish streams: species composition and environmental factors. Aquat. Bot. 66, 255–272.
Shannon, C.E., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, Urbana, IL.
Southwood, T.R.E., 1977. Habitat, templet for ecological strategies – presidential -address to British-Ecological-Society, 5 January 1977. J. Anim. Ecol. 46, 337–365.
Statzner, B., Bis, B., Doledec, S., Usseglio-Polatera, P., 2001. Perspectives for biomonitoring at large spatial scales: a unified measure for the functional composition of invertebrate communities in European running waters. Basic Appl. Ecol. 2, 73–85.
Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., Norris, R.H., 2006. Setting the expectation for the ecological condition of streams: the concept of reference condition. Ecol. Appl. 16, 1267–1276.
Tachet, H., Richoux, P., Bournaud, M., Ussegio-Polatera, P., 2006. Invertébrés d'eau douce: systématique, biologie, écologie. CNRS Editions, Paris.
Townsend, C.R., Hildrew, A.G., 1994. Species traits in relation to a habitat templet for river systems. Freshw. Biol. 31, 265–275.
Usseglio-Polatera, P., Bournaud, M., Richoux, P., Tachet, H., 2000. Biomonitoring through biological traits of benthic macroinvertebrates: how to use species trait databases? Hydrobiologia 422–423, 153–162.
van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Aquat. Ecol. 28, 117–133.
Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37, 130–137.
Yates, A.G., Bailey, R.C., 2010. Covarying patterns of macroinvertebrate and fish assemblages along natural and human activity gradients: implications for bioassessment. Hydrobiologia 637, 87–100.

P2

The relative influence of catchment, riparian corridor and local anthropogenic pressures on fish and macroinvertebrate assemblages in French rivers

Marzin, A., P. Verdonschot & D. Pont (2012)

Hydrobiologia

DOI 10.1007/s10750-012-1254-2

Hydrobiologia
DOI 10.1007/s10750-012-1254-2

The relative influence of catchment, riparian corridor, and reach-scale anthropogenic pressures on fish and macroinvertebrate assemblages in French rivers

Anahita Marzin · Piet F. M. Verdonschot · Didier Pont

Received: 9 February 2012 / Accepted: 7 July 2012

Abstract This study compares the relative influences of physiography and anthropogenic pressures on river biota at catchment, riparian corridor, and reach scales. Environmental data, catchment and riparian corridor land use, anthropogenic modifications and biological data were compiled for 301 French sites sampled from 2005 to 2008. First, relationships between anthropogenic pressures and fish and macroinvertebrate assemblages were analysed using redundancy analysis. Second, the influences of physiography and the three scales of human pressures on biological assemblages were measured using variance partitioning. Distributions of fish and macroinvertebrate taxa along the pressure gradients agreed with bio-ecological knowledge. At the reach scale, assemblage variability among the 301 French sites was related to the presence of an impoundment and to poor water quality, while at larger scales it was linked to a gradient from forest to agricultural covers. In addition, a large proportion of the explained variability in assemblage composition was related to complex interactions among factors (~40%) and to physiographic variables (~30%). Furthermore, our results highlight that catchment land use better reflects local water quality impairments than hydromorphological degradations. Finally, this study supports the idea that human pressure effects on river communities are linked at several spatial scales and must be considered jointly.

Keywords Redundancy analysis · Land use · Human pressures · Fish · Macroinvertebrate

Guest editors: C. K. Feld, A. Borja, L. Carvalho & D. Hering / Water bodies in Europe: integrative systems to assess ecological status and recovery

A. Marzin (✉) · D. Pont
Irstea, UR HBAN, HYNES (Irstea-EDF R&D), 1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony Cedex, France
e-mail: anahita.marzin@irstea.fr

P. F. M. Verdonschot
Alterra, Wageningen University and Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands

Introduction

At the river reach scale, aquatic biota respond to local environmental factors (e.g. physical habitat and water chemistry), that in turn are influenced by larger-scale pressures, such as riparian corridor condition or catchment land uses (Frissell et al., 1986; Poff, 1997; Angermeier & Winston, 1998; Bedoya et al., 2011). For instance, declines in forest cover and increases in agricultural and urban land uses are frequently related to degraded riverine habitat and biota (Allan et al., 1997; Snyder et al., 2003; Buck et al., 2004; Kroll et al., 2009). An increase in agricultural land use often results in nutrient enrichment and riffle sedimentation (Richards & Host, 1994; Snyder et al., 2003). Accordingly, the idea that the degree of river degradation can be

Published online: 24 July 2012

estimated by relative land-use cover has become common (Naiman, 1992; Allan & Flecker, 1993; Allan, 2004). Concurrently, geographic information system (GIS) technologies have considerably improved over the last few decades. Calculation of land cover reflecting human activities in a specific area is now an accessible and affordable routine task. River managers are therefore increasingly prone to replace fastidious and time-intensive field measurements of habitat and biota with money-saving and easy-to-acquire proxies for ecological condition such as land-use cover. In this context, it is of prime interest to understand how land use relates to the variability of river, biological communities in comparison to reach-scale human pressures and physiographic landscape features.

Although human impacts on river biological communities were firstly described at the reach scale (Fausch et al., 2002; Durance et al., 2006), riparian corridor- and catchment-scale effects are now well documented as well (Bis et al., 2000; Hrodey et al., 2009). For example, Roth et al. (1996) and Argent & Carline (2004) observed that fish species richness and biotic integrity decline for sub-watersheds with higher percentages of agricultural and developed land but remain relatively stable in areas with higher proportions of wetland and forest cover. Richards & Host (1994) found the same correlation for macroinvertebrate assemblages. More generally, biological integrity is positively correlated with % pasture, % tributary area, and riparian condition, and negatively correlated with % urban and farming area (e.g. Pinto et al., 2006). Nevertheless, these relationships have rarely been quantified. In addition, interaction effects between pressures at different spatial scales and physiographic factors may affect the biota. For instance, Snyder et al. (2003) demonstrated that in catchments with steeper channel slopes, the decreased contact time between riparian areas and stream channels (during which human-induced disturbances are mitigated) results in stronger urban land use impacts on fish assemblages.

Several studies have compared the ability to explain the variability in biological assemblages at these different spatial scales but results were not always consistent (Richards et al., 1997; Lammert & Allan, 1999; Nerbonne & Vondracek, 2001; Wang et al., 2003; Brazner et al., 2005; Moerke & Lamberti, 2006; Esselman & Allan, 2010; Sáli et al., 2011). Whereas some of these studies have shown that larger-scale factors are less important than reach-scale variables in explaining the variability observed in biological assemblages (Richards et al., 1997; Lammert & Allan, 1999; Wang et al., 2003; Johnson et al., 2007), others showed that local variables were weaker predictors than regional land use (Roth et al., 1996; Allan et al., 1997; Nerbonne & Vondracek, 2001; Esselman & Allan, 2010). Moreover, most of these studies did not distinguish environmental factors that were quasi-independent of human activity (here called physiographic factors) from those directly influenced by human activity (commonly called human pressure factors). Biological indicators are always affected by both natural variability and human pressures (Stoddard et al., 2006); therefore, natural variability due to physiographic differences must be considered when assessing the impact of stressors (Johnson et al., 2006; Marzin et al., 2012).

The objective of this study was to compare the relative influence of anthropogenic pressures defined at different scales (catchment land use, riparian corridor land use and reach-scale pressures) on macroinvertebrate and fish assemblages while differentiating between the influence of physiographic factors and anthropogenic pressures. Three questions were addressed: (i) How are catchment and riparian corridor land use and reach-scale pressures related? (ii) How are human pressure variables and river biological assemblage composition related? (iii) How much of the variability in fish and macroinvertebrate assemblages is explained by physiographic variability, human-induced pressures at the reach scale, riparian land use and catchment land use?

Based on the results of previous studies, we expected to observe (1) strong relationships between land use and reach-scale human pressure variables and (2) relationships between pressure variables and biological assemblage compositions. We anticipated that (3) variability in biological community composition is more closely related to physiographic factors and reach pressures than to catchment and riparian land use and that (4) complex interaction effects exist among these spatially different pressures. Finally, biological groups respond differently to human pressures (Lammert & Allan, 1999; Marzin et al., 2012), and thus we expect that (5) the results will be partly different for macroinvertebrate and fish assemblages.

The originality of this paper lies in simultaneously taking into account reach-scale human pressures and land use, and explicitly distinguishing impacts of physiography and human pressures while using a large data set covering a wide range of physiographic and pressures gradients. Our predictions were examined using French national data for 301 river sites. To define the relationship between anthropogenic pressures and biological assemblages while accounting for the effect of physiographic variability beforehand, partial redundancy analyses were conducted at three spatial scales (reach, riparian and catchment) for each biological group. Variance partitioning was used to measure the influences of physiography and the three scales of anthropogenic variables on river biological assemblages.

Materials and methods

Site sampling and biological data

As part of the European WISER project, physiographic environment, land use, reach-scale human-induced modifications and biological data were compiled from French monitoring programs for 301 sites in France (Fig. 1). The sites were sampled from 2005 to 2008 during the low-flow period using standardized protocols. Fish were sampled during national fishery surveys using electro-fishing techniques (AFNOR, 2004a). Macroinvertebrates were collected during national monitoring programs using the Indice Biologique Global Normalisé (IBGN) method (AFNOR, 2004b) and most samples were identified to the family level. Rare fish and macroinvertebrate taxa were excluded (present in fewer than 5 and 30 sites, respectively). Biological data were compiled into two taxonomic composition matrixes (i.e., abundance of each taxa with sites as rows and taxa as columns) and were subjected to the Hellinger transformation to acquire relative abundance data (Borcard et al., 2011).

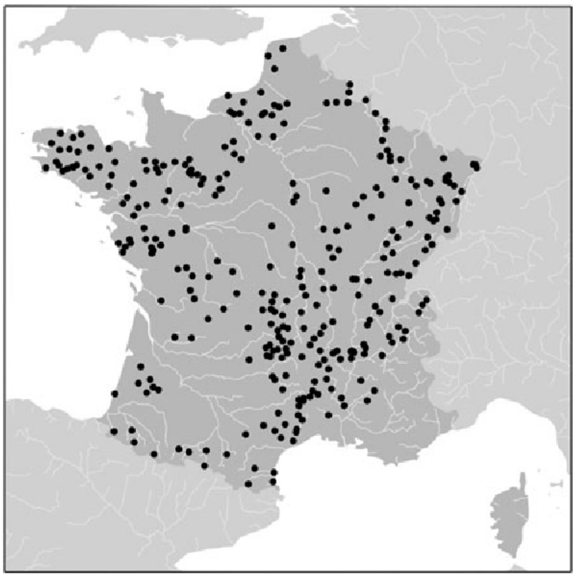


Fig. 1 Sampling sites

Physiographic data

Environmental features quasi-independent from human activity were described at the reach and catchment scales by ten variables including river habitat descriptors, such as altitude, mean annual air temperature, and air temperature range (Table 1). In addition, river type characteristics were described by distance to the river's source, upstream drainage area, type of hydrological regime, and geology. The Hill and Smith multivariate analysis (1976) was used to summarize these variables (Table 1) into three independent physiographic gradients (P1, P2, and P3, the first three axes of the analysis). Quantitative variables were log-transformed when necessary to better fulfill the normality assumption.

Land use

Land uses within the upstream drainage basin (catchment scale) and the riparian corridor (a 720-m wide buffer that extended 10,000 m upstream from each sampling site) were described by the proportion (% of the total area) of five land-use categories: farming land (FAR), artificial land (ART; developed areas, mainly urban and industry), forest and semi-natural land (FOR), wetland (WET), and water cover (WAT) (see Table 1). Land covers were calculated using CORINE Land Cover (CLC, 2000) land use maps and the ESRI ArcGIS TM 9.0 software. As proportional data, land covers were arcsine squared root transformed in order to improve normality of the distributions.

Reach-scale, human-induced modifications

At the reach scale, human-induced modifications of the river were described using 19 variables (Table 1).

Table 1 Physiographic and pressure data for 301 sites

Variables	Transformation	Modalities for qualitative variables (number of sites) and ranges for quantitative variables (median (min–max))
Physiography		
Altitude (m)[a]	Log(x)	200 (2–1520)
Stream power (kg m s^{-3})[a]	Log(x)	62920 (94–2,964,000)
Mean slope (‰)[a]	Log(x)	3.7 (0.1–82)
Catchment area (km^2)[a]	Log(x)	97 (1–13310)
Mean air temperature (°C)[a]	Log(x)	10.5 (5–15.6)
Temperature range (°C)[a]	Log(x)	24.7 (15.4–30.5)
Distance to the source (km)[a]	Log(x)	17 (0.6–372)
Hydrological regime[a]	–	Pluvial strong (121)/Pluvial moderate (165)/ Nival-glacial (27)
Geological type[a]	–	Siliceous (180)/Calcareous (133)
Ecoregions[a]	–	Alps (6)/Central highlands (7)/ Mediterranean (3)/Pyrenees (3)/ Western highlands (125)/ Western plains (169)
Reach-scale human pressures		
Presence of an impoundment	–	No (285)/Yes (28)
Barrier downstream	–	No (186)/Partial (77)/Yes (50)
Riparian vegetation modified[b]	–	No (193)/Slight (85)/Intermediate (23)/High (12)
Artificial embankment[b]	–	No (272)/Partial (23)/Yes (18)
Instream habitat modified[b]	–	No (246)/Intermediate (40)/High (27)
Channel form modified[b]	–	No (258)/Intermediate (32)/High (23)
Cross-section modified[b]	–	No (258)/Intermediate (26)/High (29)
Diked[b]	–	No (283)/Intermediate (21)/High (9)
Sedimentation[b]	–	No (177)/Slight (71)/Intermediate (47)/High (18)
Hydrological regime modified[c]	–	No (193)/Slight (70)/Intermediate (28)/High (22)
Hydropeaking[c]	–	No (285)/Yes (28)
Water abstraction[c]	–	No (195)/Slight (84)/Intermediate (12)/High (22)
Oxygen saturation (%)[d]	Log(x)	94 (46–112)
BOD5 (mg O_2/l)[d]	Log(x)	1.4 (0.5–4.5)
Nitrite (mg NO_2/l)[d]	Log(x)	0.03 (0.01–0.34)
Nitrate (mg NO_3/l)[d]	Log(x)	7.5 (0.3–74)
Ammonia (mg NH_4/l)[d]	Log(x)	0.05 (0.01–0.6)
Ortho phosphate (μg PO_4/l)[d]	Log(x)	65 (10–880)
Total phosphate (μg P/l)[d]	Log(x)	50 (10–490)
Riparian corridor land use		
% Farming	$\sqrt{(\arcsin(x))}$	62 (0–100)
% Artificial	$\sqrt{(\arcsin(x))}$	0 (0–46)
% Forest and semi-natural	$\sqrt{(\arcsin(x))}$	32 (0–100)
% Wetland	$\sqrt{(\arcsin(x))}$	0 (0–43)
% Waterland	$\sqrt{(\arcsin(x))}$	0
Catchment land use		
% Farming	$\sqrt{(\arcsin(x))}$	57 (0–100)
% Artificial	$\sqrt{(\arcsin(x))}$	1 (0–11)
% Forest and semi-natural	$\sqrt{(\arcsin(x))}$	41 (0–100)
% Wetland	$\sqrt{(\arcsin(x))}$	0 (0–6)
% Waterland	$\sqrt{(\arcsin(x))}$	0 (0–3)

Used in [a] physiographic gradient, [b] hydrological degradation gradient, [c] morphological gradient, and [d] water quality degradation gradient

These data included expert judgement on the hydrological (e.g. water abstraction, hydropeaking) and morphological modifications of river characteristics (e.g. instream habitat, channel form modifications) and median values of the water quality parameters for the 2005–2007 period (e.g. nitrate concentration). In order to reduce the correlation among variables and attend to model parsimony, five categories of pressures were defined: the presence of an impoundment (IMPOUNDMENT), the presence of a barrier downstream (BAR DOWN), water quality degradation (WQ), hydrological disturbances (HYDRO) and morphological degradations (MORPHO). For the last three pressures, the composing variables were summarized into pressure gradients using the first axis of principal component analysis (PCA) for WQ and multiple correspondence analysis (MCA) for HYDRO and MORPHO (Table 1).

Data analysis

First, to describe the relationship between land use and reach-scale human-induced modifications, Spearman (ordinal variables) and Pearson correlations (quantitative variables) were calculated among land cover types at the two scales (riparian, catchment) and among reach stressors. To control the Type I error rate over the correlations group, Bonferroni corrections were applied (Rice, 1989).

Second, partial constrained redundancy analyses (pRDAs) were used to display the patterns of the taxonomic composition of fish and invertebrate assemblages, respectively, explained by a linear model of anthropogenic pressure variables at each of the three scales, while holding the effect of physiography constant (Borcard et al., 2011):

$$\text{Biological taxonomic composition} \sim \text{Human pressure} | \text{Physiography}$$

This method allows one to account for the effect of physiography before looking at the effects of human pressures on biological communities. It is of primary importance to not confound the two signals. For each biological group, one pRDA per scale was conducted (pRDA0 = reach, pRDA1 = corridor, pRDA2 = catchment), resulting in six independent analyses.

Finally, unique and shared contributions to the explained variability of the biological assemblages were calculated for the anthropogenic explanatory variables at the three different scales and physiographic variables using variation partitioning (Borcard et al., 2011). The corridor area is a subset of the catchment area. For this reason corridor and catchment variables were never incorporated into the same analysis to avoid bias by taking into account the same information twice in the model. Thus, two analyses were conducted for each biological group: one including reach, corridor and physiographic variables (VP1) and the second including reach, catchment and physiographic variables (VP2). Adjusted redundancy statistics R^2 were calculated to provide unbiased estimates of the explained fraction of variance for all the analyses (Peres-Neto et al., 2006).

Results

Biological data

Forty-two fish species (141,211 individuals) and 88 macroinvertebrate taxa (703,135 individuals) remained after removing rare taxa. These species were represented by 15 families of fishes and 67 macroinvertebrate families, with Cyprinidae (21 species), Percidae (three species) and Salmonidae (three species) being the most diverse for fish and Elmidae (seven genera), Baetidae (three genera) and Leptoceridae (three genera) for macroinvertebrates. Five species were dominant in fish assemblages: brown trout (*Salmo trutta fario*; Linnæus, 1758), Eurasian minnow (*Phoxinus phoxinus*), bullhead (*Cottus gobio*), stone loach (*Barbatula barbatula*), and gudgeon (*Gobio gobio*) occurring at 68, 60, 56, 61, and 53% of the sites and with mean relative abundances equal to 23.9, 19.7, 13.4, 6.8, and 6.8, respectively. Five taxa prevailed in macroinvertebrate assemblages: Chironomidae, Gammaridae, genus *Baetis*, Simuliidae and Oligochaeta occurring at 100, 78, 93, 88, and 97% of the sites and with mean relative abundances equal to 17.2, 14.2, 8.9, 6.7, and 6.2, respectively.

Physiographic and reach-scale pressure gradients

The first three axes of the physiographic variable analysis accounted for 51.5% of the total variability with 26.4% explained by the first axis (*P1*), 14.6% explained by the second axis (*P2*) and 10.5%

explained by the third axis (*P3*). *P1* was related to a river's longitudinal gradient, which decreased with elevation and mean slope and increased with mean annual air temperature, catchment area, and distance to the river source. *P2* was mainly related to a decrease in stream power, while *P3* was related to geological types, from siliceous to calcareous.

The WQ, HYDRO, and MORPHO pressure gradients accounted for 55.2, 23.3, and 22.0% of the total variability, respectively, and were related to modified and degraded habitat.

Land use–reach-scale pressure correlations

Corridor and catchment FAR were related to concentrations of nitrite, nitrate, ortho- and total phosphate at downstream sites (*r*, 0.48–0.77; see Table 2). They were also positively correlated to modified riparian vegetation, instream habitat degradation and sedimentation (0.37–0.53). Catchment ART was highly correlated with concentrations of ortho- and total phosphate as well as with nitrite and nitrate (0.46–0.52). Corridor and catchment FOR were strongly negatively related to nitrite, nitrate, ortho- and total phosphate (−0.47 to −0.79). They were also negatively correlated with modifications of the hydrological regime, riparian vegetation, instream habitat degradation and sedimentation (−0.32 to −0.57). At both catchment and corridor scales, WET was not significantly correlated to reach-scale pressure parameters. Catchment WAT was related to modifications of the hydrological regime, hydropeaking and presence of an impoundment.

Relationship between human-pressure variables and biological assemblages—pRDA analyses

pRDA0 (reach scale), pRDA1 (riparian corridor) and pRDA2 (catchment) explained 28, 24, and 23% of the total inertia of the fish assemblage composition,

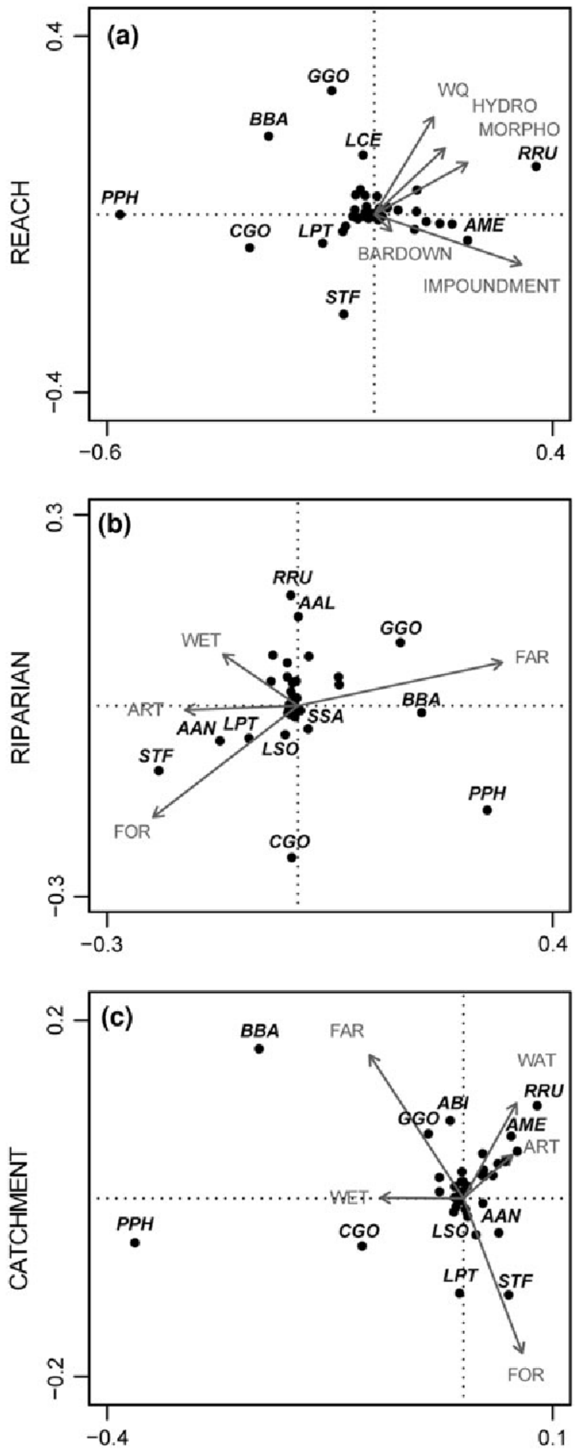
Table 2 Relationships between reach-scale and land use stressors ($|r| > 0.3$; *P* value < 0.05)

Reach pressures	Riparian corridor land use					Catchment land use				
	FAR	ART	FOR	WET	WAT	FAR	ART	FOR	WET	WAT
Impoudment	–	–	−0.31	–	–	–	0.32	–	–	0.31
Bar down	–	–	–	–	–	0.30	–	−0.31	–	–
Riparian vegetation modified	**0.43**	–	**−0.44**	–	–	0.37	0.37	−0.39	–	–
Artificial embankment	–	–	–	–	–	–	–	–	–	–
Instream habitat modified	**0.46**	–	**−0.50**	–	–	**0.40**	0.30	**−0.41**	–	–
Channel form modified	–	–	−0.30	–	–	–	–	–	–	–
Cross-section modified	0.31	–	−0.36	–	–	–	0.31	–	–	–
Diked	–	–	–	–	–	–	–	–	–	–
Sedimentation	**0.53**	–	**−0.57**	–	–	**0.49**	0.39	**−0.50**	–	–
Hydrological regime modified	0.34	0.34	−0.39	–	–	0.31	0.39	−0.32	–	0.33
Hydropeaking	–	–	–	–	–	–	–	–	–	**0.41**
Water abstraction	–	–	–	–	–	–	–	–	–	–
Oxygen saturation (%)	–	–	–	–	–	–	–	–	–	–
BOD5 (mg O_2/l)	–	–	−0.32	–	–	0.32	0.37	−0.38	–	–
Nitrite (mg NO_2/l)	**0.45**	–	**−0.52**	–	–	**0.50**	**0.50**	**−0.57**	–	–
Nitrate (mg NO_3/l)	**0.64**	–	**−0.63**	–	–	**0.77**	**0.46**	**−0.79**	–	–
Ammonia (mg NH_4/l)	–	–	–	–	–	–	0.36	–	–	–
OrthoPhosphate (µg PO_4/l)	**0.41**	0.33	**−0.47**	–	–	**0.48**	**0.52**	**−0.52**	–	–
TotalPhosphate (µg P/l)	**0.43**	0.31	**−0.51**	–	–	**0.50**	**0.52**	**−0.56**	–	–

In bold, $|r| > 0.4$

FAR farming, *ART* urban and developed, *FOR* forest, *WET* wetland, *WAT* water

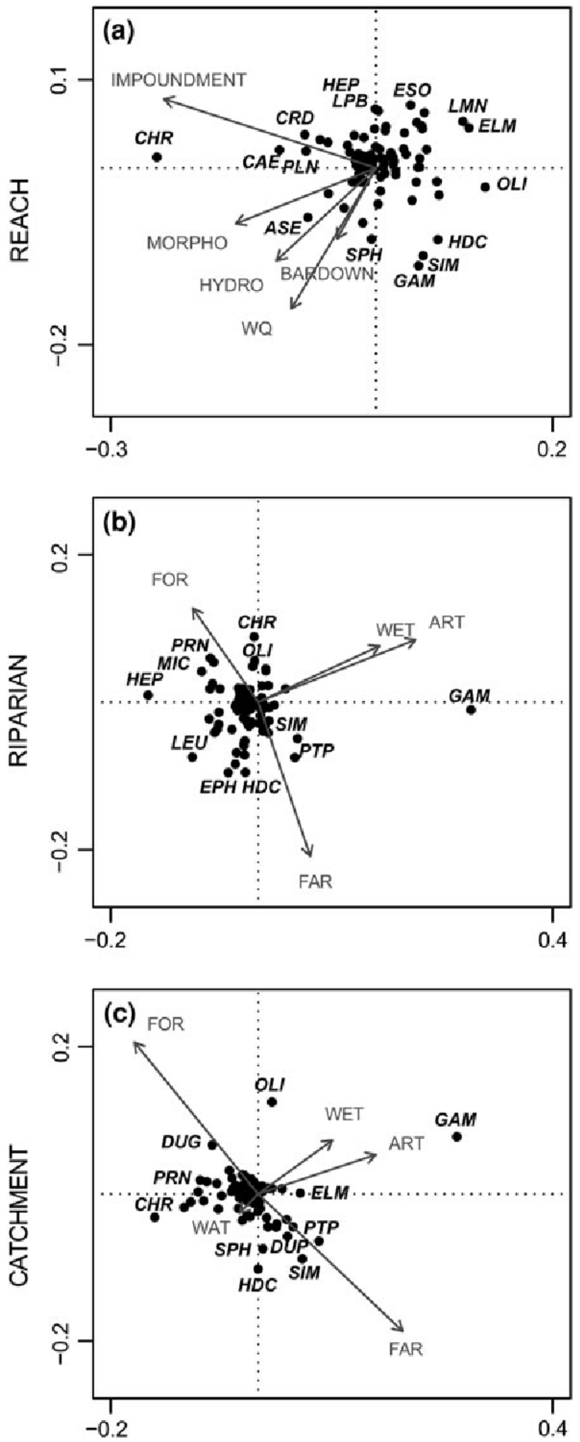
Note: Pearson correlations were computed for quantitative variables and Spearman correlation for ordinal variables



◀ **Fig. 2** Biplots from the three independent redundancy analyses (pRDA) using anthropogenic variables at three spatial scales while holding the effect of physiography constant: **a** REACH (pRDA0), **b** RIPARIAN corridor (pRDA1), and **c** CATCHMENT (pRDA2). *Arrow length* corresponds to the strength of relationships among the variables and the axes. pRDA0, pRDA1, and pRDA2 explained 28, 24, and 23% of the total inertia of the fish assemblage composition, respectively (permutation F test, $P \leq 0.001$). Three letter codes represent fish species (ABI, *Alburnoides bipunctatus*; AAL, *Alburnus alburnus*; AME, *Ameiurus melas*; AAN, *Anguilla anguilla*; BBA, *Barbatula barbatula*; CGO, *Cottus gobio*; GGO, *Gobio gobio*; LPT, *Lampetra sp.*; PPH, *Phoxinus phoxinus*; RRU, *Rutilus rutilus*; SSA, *Salmo salar*; STF, *Salmo trutta fario*; LCE, *Squalius cephalus*; LSO, *Telestes souffia*)

respectively (permutation F test, $P \leq 0.001$), suggesting a strong influence of abiotic variables on fish assemblages. Influences of pressure variables on fish assemblages were mostly reflected by the first (70, 47, and 49%, respectively; permutation F test, $P \leq 0.001$) and second axes of the pRDAs (20, 38, and 22%, respectively; permutation F test, $P \leq 0.001$). The most important explanatory variables on the first axes (the strongest correlation with the first axes) were MORPHO and impoundment for pRDA0 ($r = 0.55$ and 0.86, respectively; Fig. 2a), FAR and FOR for pRDA1 ($r = 0.66$ and -0.47, respectively; Fig. 2b) and FAR for pRDA2 ($r = -0.38$; Fig. 2c). The second axes of the pRDAs showed the strongest correlation with HYDRO and WQ for pRDA0 ($r = 0.38$ and 0.57, respectively; Fig. 2a), FOR for pRDA1 ($r = -0.36$; Fig. 2b) and FOR for pRDA2 ($r = -0.62$; Fig. 2c).

The pRDA0, pRDA1, and pRDA2 significantly explained 13, 14, and 13% of the total inertia of the macroinvertebrate assemblage composition, respectively (permutation F test, $P \leq 0.001$), suggesting a moderate influence of selected abiotic variables on macroinvertebrate assemblages. The influence of pressure variables on macroinvertebrate assemblages was mostly reflected by the first (52, 45, and 47%, respectively; permutation F test, $P \leq 0.001$) and second axes of the pRDAs (25, 28, and 21%, respectively; permutation F test, $P \leq 0.001$). The most important explanatory variables on the first axes were MORPHO and IMPOUNDMENT for pRDA0 ($r = -0.57$ and -0.86, respectively; Fig. 3a), ART and WET for pRDA1 ($r = 0.64$ and 0.50, respectively; Fig. 3b) and FAR, FOR and ART for pRDA2 ($r = 0.54$, -0.46, and 0.44, respectively; Fig. 3c). The second axes of the pRDAs showed the strongest correlations with HYDRO and WQ for pRDA0 ($r = -0.38$ and -0.57, respectively; Fig. 3a), FAR



◀ **Fig. 3** Biplots from the three independent redundancy analyses (pRDA) using anthropogenic variables at three spatial scales while holding the effect of physiography constant: **a** REACH (pRDA0), **b** RIPARIAN corridor (pRDA1), and **c** CATCHMENT (pRDA2). *Arrow length* corresponds to the strength of relationships among the variables and the axes. pRDA0, pRDA1, and pRDA2 explained 13, 14, and 13% of the total inertia of the macroinvertebrate assemblage composition, respectively (permutation F test, $P \leq 0.001$). Three letter codes represent macroinvertebrate taxa (ASE, *Asellidae Gen.* sp.; CAE, *Caenis* sp.; CHR, Chironomidae Gen. sp.; CRD, *Corixidae Gen.* sp.; DUG, *Dugesiidae Gen.* sp.; DUP, *Dupophilus* sp. *Ad.*; ELM, *Elmis* sp. *Ad.*; EPH, *Ephemera* sp.; ESO, *Esolus* sp. *Ad.*; GAM, *Gammaridae Gen.* sp.; HEP, *Heptageniidae Gen.* sp.; HDC, *Hydropsyche* sp.; LPB, *Leptophlebiidae Gen.* sp.; LEU, *Leuctridae Gen.* sp.; LMN, *Limnius* sp. *Ad.*; MIC, *Micrasema* sp.; OLI, *Oligochaeta Gen.* sp.; PLN, *Planorbidae Gen.* sp.; PTP, *Potamopyrgus* sp.; PRN, *Protonemura* sp.; SIM, *Simuliidae Gen.* sp.; SPH, *Sphaerium* sp.)

and FOR for pRDA1 ($r = -0.63$ and 0.38, respectively; Fig. 3b) and FOR and FAR for pRDA2 ($r = 0.56$ and -0.51, respectively; Fig. 3c).

Amount of variability in fish and macroinvertebrate assemblages explained by physiography and human pressures at different scales

In the four variance partitioning analyses, all types of variables significantly contributed to the explained variation (P value < 0.001). In the analysis including reach, corridor and environmental variables (VP1), 30% of the total variability in fish assemblage composition was explained. Solely riparian land use, solely reach pressures and solely physiographic factors accounted for 6, 20 and 31% of the explained variability, respectively, while 43% was explained by the shared factors (RIP + PHY = 10%, REACH + -PHY = 5%, RIP + REACH = 3%, ALL = 25%; Fig. 4a). In comparison, 29% of the variability in fish species abundance was explained for the analysis including reach, catchment and environmental variables (VP2). Pure catchment land use, pure reach pressures, and pure physiographic factors contributed 5, 21, and 27% of the explained variability, respectively, while 47% was explained by the shared factors (CATCH + PHY = 14%, REACH + PHY = 2%, CATCH + REACH = 3%, ALL = 28%; Fig. 4b).

2VP1 and VP2 each explained 15% of the total variation in macroinvertebrate taxonomic composition. In the first analysis, solely riparian land use, solely reach pressures and solely physiographic factors accounted for 11, 10, and 38% of the explained variability, respectively, while 41% was explained by

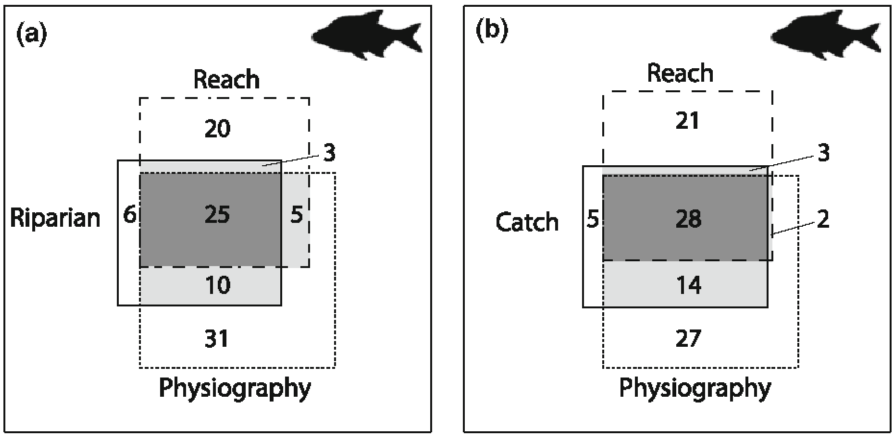


Fig. 4 Venn diagrams representing the results of the two RDA analyses VP1 (**a**) and VP2 (**b**): amount of variation in fish species abundance explained by physiographic factors (*dotted line square*), anthropogenic factors at the reach (*dashed line square*), **a** riparian corridor (*solid line square*) and **b** catchment scales (*solid line square*) and their interactions. Each area is proportional to the share of the inertia explained by the single factor (*white area*) or its interactions with other corresponding factors (*grey areas*). *Numbers* correspond to the percentage of the explained variation associated with each variable type. Thirty percent of the total variance was explained for VP1 and 29% for VP2

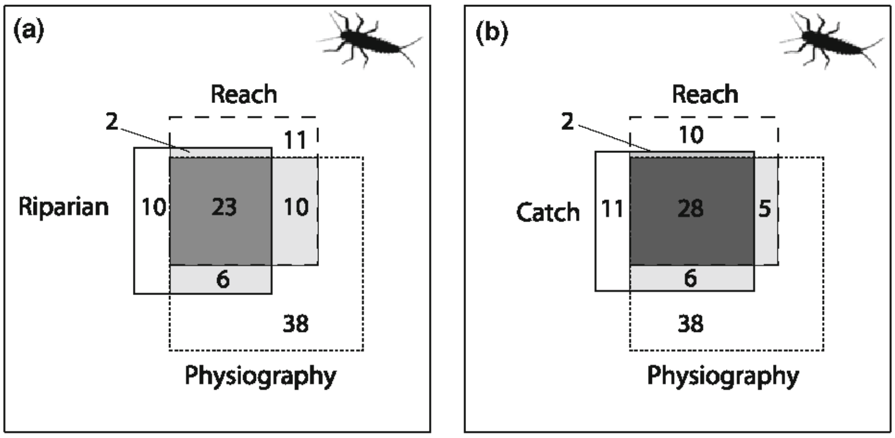


Fig. 5 Venn diagrams representing the results of the two RDA analyses VP1 (**a**) and VP2 (**b**): amount of variation in macroinvertebrate taxa abundance explained by physiographic factors (*dotted line square*), anthropogenic factors at the reach (*dashed line square*), **a** riparian corridor (*solid line square*) and **b** catchment scales (*solid line square*) and their interactions. Each area is proportional to the share of inertia explained by the single factor (*white area*) or its interactions with other corresponding factors (*grey areas*). *Numbers* correspond to the percentage of the explained variation associated with each variable type. For both analyses, 15% of the total variance was explained

the shared factors (RIP + PHY = 6%, REACH + - PHY = 10%, RIP + REACH = 2%, ALL = 23%; Fig. 5a). In the second, solely catchment land use, solely reach pressures and solely physiographic factors contributed 10, 11, and 38% of the explained variability, respectively, while 41% was explained by the shared factors (CATCH + PHY = 6%, REACH + PHY = 5%, CATCH + REACH = 2%, ALL = 28%; Fig. 5b).

Discussion

Land use as a proxy for reach-scale anthropogenic pressure variables

In our data sets, water quality parameters were on average more strongly correlated to land-use cover than hydro-morphological parameters. These results imply that when considering land cover as a proxy for reach-scale habitat or water quality degradation, water quality problems will be better represented than local habitat and hydro-morphological problems. However, some differences were observed between catchment and corridor land use. Local water quality parameters were better correlated to catchment than to corridor land cover, and riparian land cover was better correlated to hydro-morphological degradations. These findings are in accordance with those of several previous studies (Richards et al., 1996; Sliva & Williams, 2001), suggesting that catchment land covers are potential candidate proxies for local water quality parameters and corridor land covers are potential predictors of reach-scale habitat and hydro-morphological parameters.

These patterns were confirmed when examining the correlation of the different land cover categories. At the catchment scale, percentages of forested area showed the strongest negative correlations to nutrient enrichment (nitrite, nitrate, ortho- and total phosphate). At the corridor scale, percentages of forested area showed the strongest negative correlations to degradation of the riparian vegetation, instream habitats and sedimentation. On the other hand, at the catchment scale, the percentage of agricultural land uses presented the strongest positive correlations to nitrate and nitrite, and the percentage of artificial land use was strongly related to concentrations in phosphate. These relationships could be explained by the nutrient inputs from farming areas and domestic sewage and industrial waste discharges in artificial areas. The percentage of agricultural land use in riparian corridors was strongly correlated to degradation of the riparian vegetation, instream habitat, and sedimentation. The percentages of water cover in the upstream catchment presented the strongest relationship to the presence of impoundment and hydrological degradation. By contrast, wetland cover was not significantly correlated to reach-scale parameters. These variables may contribute new information on the river environment in addition to reach scale variables.

Relationship between anthropogenic pressures and biological assemblage composition

Our objective was to estimate the influence of human-induced pressure variables at different scales after having removed the variability related to physiography. The proportion of the total inertia of assemblage composition explained by the analyses was lower for macroinvertebrates than for fish. Eighty-eight taxa were used to describe the macroinvertebrate assemblages versus only 51 for fish. This difference in explained inertia could be due to greater composition complexity for the macroinvertebrate assemblages. In addition, whereas taxonomic resolution for fish was the species, macroinvertebrate were mostly identified at genus or family level. Consequently, physiography and human pressure impacts on macroinvertebrates might be partly masked by pooling several species with different preferences and responses to pressures into a single family. Furthermore, although abiotic factors explained a significant proportion of the variability in biological assemblages, the abiotic factors chosen might be more relevant to fish assemblages than to macroinvertebrate assemblages. Considering these issues, macroinvertebrate results should be considered with caution.

The pRDA results were consistent at the three scales and with numerous previous studies demonstrating the important role played by human-induced pressures on the species composition of riverine assemblages (Richards et al., 1996; Moerke & Lamberti, 2006). In our study, the presence of an impoundment emerged as the main human pressure factor shaping the fish and macroinvertebrate assemblages at the reach scale, followed by water quality and morphological pressure gradients. Indeed, the presence of an impoundment is known to have major impacts on river assemblages (Baxter, 1977; Ward & Stanford, 1979; Tiemann et al., 2004) because it changes river functioning considerably by shifting from a free-flowing water system to stagnant water. Our results suggest that taxa that are pollution resistant, eurythermic, limnophilic and those preferring a slow velocity, such as roach, black bullhead (Kottelat & Freyhof, 2007) and Chironomidae (Tachet et al., 2006), are positively influenced by the presence of an impoundment and reach-scale morphological degradations. By contrast, taxa preferring relatively cold running waters, such as Gammaridae, Simuliidae or the *Hydropsyche* genus (Tachet et al., 2006) and intolerant to low dissolved oxygen concentrations, such as Eurasian minnow, brown trout, bullhead and lamprey (Kottelat & Freyhof, 2007) are negatively related to impoundment. Moreover, fish species that are more tolerant, rheophilic, and those preferring to spawn in running water, such as gudgeon or stone loach (Kottelat & Freyhof, 2007), seem to be negatively impacted by impoundment and positively impacted by poor water quality.

At broader scales (corridor and catchment), fish and macroinvertebrate assemblages appear to be strongly influenced by a common gradient from forest to agricultural land use. For instance, stenotherm-intolerant fish species preferring to spawn in running water such as brown trout, bullhead and lamprey seem linked to forested land cover, probably in relation with the dominance of small streams. More tolerant taxa, such as gudgeon, stone loach and minnow for fish and Simuliidae, Leuctridae, *Hydropsyche*, *Potamopyrgus*, and *Dupophilus* genera for macroinvertebrates appear to be related to agricultural land cover.

In addition, an increase in corridor artificial and wetland cover appears to be another important gradient influencing macroinvertebrate assemblage composition, particularly the abundance of Gammaridae. Oligochaeta were negatively correlated to reach-scale pressures and farming land use and positively correlated to catchment forest cover and artificial land use. Oligochaeta are often recognized as species tolerant to water quality pollution and eutrophication (Lafont et al., 1996; Verdonschot, 2006). However, some species present more intolerant traits and are generally present in good ecological class rivers as Naididae species (Tachet et al., 2006; Verdonschot, 2006). It is likely that in this study, species grouped under the Oligochaeta family were mostly represented by such pollution-intolerant species.

Amount of fish and macroinvertebrate assemblage variability explained by physiography and human pressures at different scales

As expected, given previous studies showing that local and larger-scale physiographic variables shape the river communities (Logez et al., 2010), physiographic variables accounted for a large proportion of the explained variability of assemblage composition among sites. For both macroinvertebrates and fish, they explained approximately one-third of the variability in assemblage composition. These results strengthen the idea that physiographic variability is a key parameter explaining river assemblage composition diversity and should be considered beforehand when looking at the effects of human-induced pressures on river ecological quality. In comparison with the classical typology retained in the Water Framework Directive (European Union, 2000), our approach has the same objective, but directly integrates the physiographic variables into the analysis, avoiding splitting the river continuum in different river types. A large share of the explained variability in assemblage composition was related to shared-factor effects (around 40% of the explained variability). For both fish and macroinvertebrate assemblages, more than half of these effects were related to the shared effects of physiographic and human pressure variables at all scales, while about one-third of these effects strictly involved the shared effects of human pressures and physiography and a few percent involved the shared effects of reach-scale pressures and land uses. These conclusions are in agreement with those of Allan (2004), confirming the influence of factors belonging to a wide range of spatial scales on lotic ecosystems and a difficulty distinguishing the roles of near-stream versus larger spatial scales. Such complex effects illustrate why establishing a simple pressure–impact relationship for fish and macroinvertebrates in rivers is so challenging, because single pressure effects are generally difficult to distinguish. Consequently, in the common case of multi-impacted rivers, it will be very hard to inform water managers about the main pressure disturbing a river's ecological status. Human pressures should be considered at multiple spatial scales to show the diverse range of effects of human activity on river communities.

In addition, these results show the existence of the joint effects of physiography and human pressure variables on macroinvertebrate and fish assemblages. Subsequently, these two effects cannot be properly distinguished. Wang et al. (2006) showed that in undisturbed catchments, fish assemblages were predominantly influenced by local physiographic factors, but as disturbance increased in catchments and riparian areas, the relative importance of local factors declined and that of catchment increased. We advocate that in future studies, these interaction effects should be taken into account to analyse the effects strictly due to human pressures on river biological communities. Furthermore, in this study, the physiographic and human pressure variables explained less than half of the total variability in assemblage compositions. Finer variables such as the sediment size distribution or streamflow might have improved the analyses if they had been available for such a large data set.

The results concerning the relative influences of anthropogenic pressures were different for macroinvertebrate and fish assemblages. Fish assemblage composition appears to be more sensitive to reach-scale anthropogenic pressures than to corridor and catchment-scale land use. By contrast, land use and reach-scale variables contribute identically to determining macroinvertebrate assemblage composition. These results can be surprising given that small, sedentary organisms, such as macroinvertebrates, are generally expected to be more sensitive to local pressure while larger, migratory organisms, such as fish, are expected to be affected by pressures acting at a larger scale. Nevertheless, as previously mentioned, the response of macroinvertebrate assemblages to small scale structure might be masked by the coarse taxonomic resolution in this study. In addition, local-scale pressures were mostly described at the reach scale and only a few at the habitat scale. Furthermore, our results support that land-use variables mainly reflect water quality degradations of the reach and upstream drainage basin area. Macroinvertebrate assemblages are generally more sensitive than fish assemblages to water quality degradation (Hering et al., 2006; Johnson et al., 2006; Justus et al., 2010; Marzin et al., 2012). Since fish and macroinvertebrate assemblages are influenced differently by human pressure variables at different scales, it appears likely that these groups provide complementary information on the river's ecological status. These results are in accordance with Flinders et al. (2008), reinforcing the idea that the use of multiple biological groups may be appropriate when developing monitoring programs.

Conclusion

Our findings support the notion that fish assemblages are more responsive to reach-scale anthropogenic pressures than to land use, whereas macroinvertebrate assemblages present comparable sensitivity to land use and reach-scale variables. Given their different responses, the use of multiple biological groups may be appropriate to monitor river ecosystems. Our results suggest that reach-scale monitoring is essential to describe the response of river ecosystems to human-induced pressures. Indeed, although the proportion of forested area in the upstream catchment and riparian corridor appears to be a useful surrogate for overall water quality degradation, we advocate that land-use information should always be combined with information on local-scale pressures, as it poorly represents reach-scale hydro-morphological impacts. More generally, this study shows that it is not possible to limit the description of human pressures impacting river macroinvertebrate and fish assemblages to one spatial scale, especially just catchment land use, without losing important information. To isolate or determine the main pressure affecting a river ecosystem will often be complicated and will need to consider interaction effects. Therefore, the variability of physiography and its interactions with human pressures at different spatial scales should always be considered when analysing the effect of human-induced pressures on river ecological quality. Future advanced research on the complex effects of the interactions among human pressures and between pressures and the environment would advance both management and understanding of river ecosystems.

Acknowledgments This paper is part of the WISER project (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (Contract No. 226273), www.wiser.eu. This work has been supported through a collaboration between Irstea and the French Electric Company (EDF): "HYNES".

References

AFNOR, 2004a. Qualité de l'eau—Détermination de l'indice poissons rivières (IPR). Association française de normalisation. Norme homologuée T90-344.

AFNOR, 2004b. Qualité écologique des milieux aquatiques. Qualité de l'eau—Détermination de l'indice biologique global normalisé (IBGN). Association française de normalisation. Norme homologuée T90-350.

Allan, J. D., 2004. Influence of land use and landscape setting on the ecological status of rivers. Limnetica 23: 187–198.

Allan, J. D. & A. S. Flecker, 1993. Biodiversity conservation in running waters. BioScience 43: 32–43.

Allan, J. D., D. L. Erickson & J. Fay, 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37: 149–161.

Angermeier, P. L. & M. R. Winston, 1998. Local vs. regional influences on local diversity in stream fish communities of Virginia. Ecology 79: 911–927.

Argent, D. G. & R. F. Carline, 2004. Fish assemblage changes in relation to watershed landuse disturbance. Aquatic Ecosystem Health and Management 7: 101–114.

Baxter, R. M., 1977. Environmental effects of dams and impoundments. Annual Review of Ecology and Systematics 8: 255–283.

Bedoya, D., E. S. Manolakos & V. Novotny, 2011. Characterization of biological responses under different environmental conditions: a hierarchical modeling approach. Ecological Modelling 222: 532–545.

Bis, B., A. Zdanowicz & M. Zalewski, 2000. Effects of catchment properties on hydrochemistry, habitat complexity and invertebrate community structure in a lowland river. Hydrobiologia 422–423: 369–387.

Borcard, D., F. Gillet & P. Legendre, 2011. Numerical Ecology with R. Use R! Series. Springer, New York.

Brazner, J. C., D. K. Tanner, N. E. Detenbeck, S. L. Batterman, S. L. Stark, L. A. Jagger & V. M. Snarski, 2005. Regional, watershed, and site-specific environmental influences on fish assemblage structure and function in western Lake Superior tributaries. Canadian Journal of Fisheries and Aquatic Sciences 62: 1254–1270.

Buck, O., D. K. Niyogi & C. R. Townsend, 2004. Scale-dependence of land use effects on water quality of streams in agricultural catchments. Environmental Pollution 130: 287–299.

Durance, I., C. Lepichon & S. J. Ormerod, 2006. Recognizing the importance of scale in the ecology and management of riverine fish. River Research and Applications 22: 1143–1152.

Esselman, P. C. & J. D. Allan, 2010. Relative influences of catchment- and reach-scale abiotic factors on freshwater fish communities in rivers of northeastern Mesoamerica. Ecology of Freshwater Fish 19: 439–454.

European Union, 2000. Directive 2000/60/EC. Establishing a Framework for Community Action in the Field of Water Policy. European Commission PE-CONS 3639/1/100 Rev 1, Luxemburg.

Fausch, K. D., C. E. Torgersen, C. V. Baxter & H. W. Li, 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. Bioscience 52: 483–498.

Flinders, C. A., R. J. Horwitz & T. Belton, 2008. Relationship of fish and macroinvertebrate communities in the mid-Atlantic uplands: implications for integrated assessments. Ecological Indicators 8: 588–598.

Frissell, C. A., W. J. Liss, C. E. Warren & M. D. Hurley, 1986. A hierarchical framework for stream habitat classification—viewing streams in a watershed context. Environmental Management 10: 199–214.

Hering, D., R. K. Johnson, S. Kramm, S. Schmutz, K. Szoszkiewicz & P. F. M. Verdonschot, 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshwater Biology 51: 1757–1785.

Hill, M. O. & A. J. E. Smith, 1976. Principal component analysis of taxonomic data with multi-state discrete characters. Taxon 25: 249–255.

Hrodey, P. J., T. M. Sutton, E. A. Frimpong & T. P. Simon, 2009. Land-use impacts on watershed health and integrity in Indiana warmwater streams. American Midland Naturalist 161: 76–95.

Johnson, R. K., D. Hering, M. T. Furse & R. T. Clarke, 2006. Detection of ecological change using multiple organism groups: metrics and uncertainty. Hydrobiologia 566: 115–137.

Johnson, R. K., M. T. Furse, D. Hering & L. Sandin, 2007. Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programmes. Freshwater Biology 52: 939–958.

Justus, B. G., J. C. Petersen, S. R. Femmer, J. V. Davis & J. E. Wallace, 2010. A comparison of algal, macroinvertebrate, and fish assemblage indices for assessing low-level nutrient enrichment in wadeable Ozark streams. Ecological Indicators 10: 627–638.

Kottelat, M. & J. Freyhof, 2007. Handbook of European Freshwater Fishes. Kottelat, Cornol.

Kroll, S. A., C. N. Llacer, M. De La Cruz Cano & J. De Las Heras, 2009. The influence of land use on water quality and macroinvertebrate biotic indices in rivers within Castilla-La Mancha (Spain). Limnetica 28: 203–214.

Lafont, M., J. C. Camus & A. Rosso, 1996. Superficial and hyporheic oligochaete communities as indicators of pollution and water exchange in the River Moselle, France. Hydrobiologia 334: 147–155.

Lammert, M. & J. D. Allan, 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23: 257–270.

Linnæus, C., 1758. Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata.

Logez, M., D. Pont & M. T. Ferreira, 2010. Do Iberian and European fish faunas exhibit convergent functional structure along environmental gradients? Journal of the North American Benthological Society 29: 1310–1323.

Marzin, A., V. Archaimbault, J. Belliard, C. Chauvin, F. Delmas & D. Pont, 2012. Ecological assessment of running waters: do macrophytes, macroinvertebrates, diatoms and fish show similar responses to human pressures? Ecological Indicators 23: 56–65.

Moerke, A. H. & G. A. Lamberti, 2006. Scale-dependent influences on water quality, habitat, and fish communities in streams of the Kalamazoo River Basin, Michigan (USA). Aquatic Sciences 68: 193–205.

Naiman, R. J., 1992. New perspectives for watershed management: balancing long-term sustainability with cumulative environmental change. In Naiman, R. J. (ed.), Watershed Management: Balancing Sustainability and Environmental Change. Springer, NewYork.

Nerbonne, B. A. & B. Vondracek, 2001. Effects of local land use on physical habitat, benthic macroinvertebrates, and fish in the Whitewater River, Minnesota, USA. Environmental Management 28: 87–99.

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

Pinto, B. C. T., F. G. Araujo & R. M. Hughes, 2006. Effects of landscape and riparian condition on a fish index of biotic integrity in a large southeastern Brazil river. Hydrobiologia 556: 69–83.

Poff, N. L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16: 391–409.

Rice, W. R., 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.

Richards, C. & G. Host, 1994. Examining land use influences on stream habitats and macroinvertebrates: a GIS approach. JAWRA Journal of the American Water Resources Association 30: 729–738.

Richards, C., L. B. Johnson & G. E. Host, 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53: 295–311.

Richards, C., R. J. Haro, L. B. Johnson & G. E. Host, 1997. Catchment and reach-scale properties as indicators of macroinvertebrate species traits. Freshwater Biology 37: 219–230.

Roth, N. E., J. D. Allan & D. L. Erickson, 1996. Landscape influences on stream biotic integrity assessed at multiple spatial scales. Landscape Ecology 11: 141–156.

Sáli, P., P. Takács, I. Kiss, P. Bíró & T. Eros, 2011. The relative influence of spatial context and catchment- and site-scale environmental factors on stream fish assemblages in a human-modified landscape. Ecology of Freshwater Fish 20: 251–262.

Sliva, L. & D. D. Williams, 2001. Buffer zone versus whole catchment approaches to studying land use impact on river water quality. Water Research 35: 3462–3472.

Snyder, C. D., J. A. Young, R. Villella & D. P. Lemarie, 2003. Influences of upland and riparian land use patterns on stream biotic integrity. Landscape Ecology 18: 647–664.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Jonhson & R. H. Norris, 2006. Setting the expectation for the ecological condition of streams: the concept of reference condition. Ecological Applications 16: 1267–1276.

Tachet, H., P. Richoux, M. Bournaud & P. Ussegio-Polatera, 2006. Invertébrés d'eau douce: systématique, biologie, écologie. CNRS Editions, Paris.

Tiemann, J. S., D. P. Gillette, M. L. Wildhaber & D. R. Edds, 2004. Effects of lowhead dams on riffle-dwelling fishes and macroinvertebrates in a midwestern river. Transactions of the American Fisheries Society 133: 705–717.

Verdonschot, P., 2006. Beyond masses and blooms: the indicative value of Oligochaetes. Hydrobiologia 564: 127–142.

Wang, L. Z., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela & P. M. Stewart, 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences 60: 491–505.

Wang, L., P. W. Seelbach & R. M. Hughes, 2006. Introduction to landscape influences on stream habitats and biological assemblages. American Fisheries Society Symposium 48: 1–23.

Ward, J. V. & J. A. Stanford, 1979. The Ecology of Regulated Streams. Plenum Press, New York.

P3

Uncertainty associated with river health assessment in a varying environment: the case of a predictive fish-based index in France.

Marzin, A., J. Belliard, O. Delaigue, M. Logez & D. Pont.

In revision for resubmission to the *Canadian Journal of Fisheries and Aquatic Sciences*

Abstract

Sensitive biological measures of river ecosystem quality are needed to assess, maintain or restore ecological conditions of European water bodies. Since our understanding of these complex systems is imperfect, decision-making requires recognizing uncertainty. A new predictive multimetric index, IPR+, based on fish functional traits was developed for the French rivers using the reference condition approach. Information on fish assemblage structure, local environment and human-induced disturbances of 1654 French river sites was used. Among the 228 potential metrics tested, only 11 were retained for the index computation. IPR+ is sensitive to overall, hydrological, morphological and water quality degradations. A Bayesian framework was used to predict theoretical metric values in absence of pressure and to estimate the uncertainty associated with these predictions. This enabled to compute the uncertainty associated with index score and to estimate the confidence associated with the evaluation of site ecological conditions. This new methodology could be used to develop bioindication tools using different biological groups and extended to other areas.

Résumé

Des mesures biologiques sensibles à la qualité des écosystèmes sont nécessaires pour évaluer, maintenir et restaurer les conditions écologiques des masses d'eau européennes. Notre connaissance de ces systèmes complexes est imparfaite et par conséquent la prise de décision nécessite la reconnaissance de l'incertitude. Un nouvel indice multi-métrique et prédictif, IPR+, a été développé pour les cours d'eau français. Il est basé sur les traits fonctionnels des poissons et sur l'approche par condition de référence. La structure des peuplements piscicoles, l'environnement local et les perturbations induites par les activités anthropiques de 1654 tronçons ont été utilisés. Parmi les 228 métriques testées, 11 métriques sensibles aux dégradations globales, hydrologiques, morphologiques et de la qualité de l’eau ont été sélectionnées pour le calcul de l'indice IPR+. Une approche bayésienne a été utilisée pour prédire les valeurs théoriques des métriques en absence de pressions et pour estimer l'incertitude autour de ces prédictions. En propageant l’incertitude prédictive des métriques tout au long du calcul de l’indice, l'incertitude des notes de l'indice ainsi que la confiance dans la classification de l'état écologique des sites ont été estimés. Cette méthodologie relativement générique pourrait être étendue à d'autres régions du monde et utilisée pour développer des indicateurs biologiques utilisant différents groupes biologiques.

Introduction

Since the implementation of the European Water Framework Directive (WFD), the EU member states must assess, maintain and restore water bodies with regard to their ecological conditions (European Union (EC) 2000). In addition to water quality, river monitoring must integrate information on biological communities to quantify anthropogenic impacts on river functioning. WFD recommends using the Reference Condition Approach (Bailey et al. 1998) to assess stream ecological conditions. Bioindicator values should be obtained by transforming, through complex processes, deviations between observed and theoretical metric values observed in reference conditions (i.e. absence of significant human disturbance). Several bioindicators (e.g. McCormick et al. 2001) were derived from the Index of Biotic Integrity (IBI; Karr 1981) . The IBI was the first index integrating several metrics based on fish functional guilds to describe assemblage (e.g. trophy or tolerance to degradation). More recently, multi-metric indices such as the European Fish Index (Pont et al. 2006, 2007) accounted for the natural variability of assemblage descriptors. They used predictive methods to estimate reference conditions (Oberdorff et al. 2002; Pont et al. 2006, 2007; Logez and Pont 2011), which therefore could be applied at a larger scale. Whatever method is applied, the accuracy of reference condition estimation remains a critical point.

Multiple sources of uncertainty have repercussions on bioindicator scores and thus could affect the final diagnostics and result in risk, important for water managers (Funtowicz and Ravetz 1993; Clarke and Hering 2006). Few studies have measured these uncertainties. Most focused on errors related to monitoring and sampling strategy (Clarke et al. 1996; Ostermiller and Hawkins 2004; Bady et al. 2005; Staniszewski et al. 2006). Evaluation of the uncertainty related to the definition of the reference condition appears essential to estimate the uncertainty around an index value and the constituent metrics. Risk and uncertainty are fundamentally reflected by the spread of a probability distribution, i.e. what are the probability distributions of an index value for a given site's environmental conditions, in absence of human pressure?

Since the early 1980s, considerable advances in statistical theory and computing technology have facilitated the development of Bayesian statistical methods and their application to complex natural resource management problems (Brooks 2003). The explicit use of probability for quantifying uncertainty in inferences based on statistical data analysis (Gelman et al. 2004) make Bayesian methods a practical and efficient method to estimate uncertainty.

Based on these considerations, the aim of this study was to develop a predictive multi-metric index based on fish functional guilds for French rivers (IPR+), delivering an index score associated with a confidence measure for decision-makers. IPR+ is original in that it: (i) defines reference conditions from objective criteria based on pressure levels, (ii) controls the natural variability of the metric to ensure that metric deviation measures the effect of anthropogenic disturbances, (iii) estimates the uncertainty associated with metric values within a Bayesian framework and (iv) propagates uncertainty throughout the computation process to estimate final index uncertainty.

Besides the description and discussion of this new methodology, this paper evaluates (1) the capacity of IPR+ to quantify the impact of different types of human disturbances on local fish communities, regardless of the natural environmental conditions (abiotic landscape features) and (2) the variations in the predictive uncertainty of the IPR+ index with human pressure levels and natural environmental conditions.

Methods

Site selection and classification of human-induced disturbances

Data were obtained from national fisheries surveys conducted between 1998 and 2007 (Poulet et al. 2011). A total of 1654 sites including 73 species were sampled once using electric fishing (Fig. 1; CON dataset). To homogenize the sampling effort, only fish collected during the first pass were considered. Rare species were removed from the samples ($N<2$ for samples with fewer than 200 fish caught, $N<3$ for other samples).

Anthropogenic alterations of rivers were assessed considering stream morphology (e.g. channel form modifications), hydrology (e.g. hydropeaking), water quality, connectivity (presence of a barrier downstream), presence of an upstream lake and navigation (Table 1). Based on this alteration evaluation, two subsets of reference sites (REF=266 sites) and calibration sites (CAL=278 sites) were defined as not or slightly impacted sites, with more than 30 fishes sampled and a fished area greater than 100 m^2. To cover the largest environmental gradients in the modelling process, the criteria used to define CAL sites were more severe upstream than downstream following the classical Huet zonation (trout zone, grayling zone, barbel zone and bream zone; Huet 1954). In contrast, unique criteria were used for REF sites (Table 1) along the longitudinal gradient.

Global human pressures (P_{Global}), water quality degradation (P_{WQ}), hydrological degradation (P_H) and morphological degradation indexes (P_M) were summarized by the first axes of multiple correspondence analyses (MCA; Tenenhaus and Young 1985). They accounted for 9.3%, 17%, 14% and 20%, respectively, of the inertia and were related to an increase in disturbance. Four classes of disturbance level ranging from one (not or slightly impacted sites) to four (heavily impacted sites) were defined by a K-means algorithm (Hartigan and Wong 1979) and four subsets of strongly impaired sites were considered: P4 (*N*=357), WQ4 (*N*=128), H4 (*N*=149) and M4 (*N*=140). $P_{WQ,}$ P_H and PM were computed on sub-datasets presenting a single type of pressure (454, 786 and 850 sites, respectively).

Environmental variables

Six environmental descriptors known to influence fish assemblages (Pont et al. 2005; Logez et al. 2012b) were chosen so as to minimize correlations between descriptors: catchment area (CA, 0.7–110,248 km²), stream power (POW, 39–114,400,000 $kg.m.s^{-3}$), mean annual air temperature (AT, 6–17°C), mean annual air temperature amplitude (ATA, 9–20°C), catchment geological type (GT_S = siliceous, 738 sites; GT_C = calcareous, 916 sites) and hydrological regime (H_PS = Pluvial strong, 499 sites; H_PM = Pluvial moderate, 959 sites; H_NG = Nival-Glacial, 196 sites). They were either measured in the field or derived from geographical information systems. Climatic data were provided by the French Meteorological Institute (Vidal et al. 2010) and averaged for the 10 years previous to sampling. These variables were assumed to be not or slightly modified by local anthropogenic activities. Due to the skewness of their distributions, CA and POW were log-transformed.

Potential metrics

Twelve biological and ecological traits were considered according to previous classifications of European fish traits at the species level with regard to reproduction, trophic position, habitat preference, sensitivity to water quality, habitat alteration, and migratory behaviour (Noble et al. 2007; Logez et al. 2012a; see Table S1 in Supplementary material S1[1]). Each species was assigned to one of the different categories of a trait (12 traits, 37 categories). Each trait was considered either in absolute or relative density (N, N%), number of species (S, S%) or biomass (B, B%) leading to 222 potential metrics. To account for WFD requirements, six metrics based on trout young of the year (T0+) were tested as age-class metrics (Logez and

[1] French fish biological and ecological functional traits

Pont 2011). The selected T0+ was calculated and will be applicable only for sites belonging to trout or grayling zones and sampled between April and December. The 228 potential metrics were log-transformed (log(X+1)).

Metric modelling and selection

To predict metric values in the reference condition, metrics were modelled (Generalized Linear Model) as a function of the six descriptors in absence of significant human disturbance (CAL dataset). Poisson regressions were chosen to model richness metrics (S) and negative binomial models were used for abundance (N) and biomass (B) metrics. An offset parameter was added (log(S_T), log(N_T) and log(B_T) (total richness, abundance and biomass, respectively) for all the relative metrics (S%, N% and B%) (McCullagh and Nelder 1989). An Akaike information criterion stepwise selection (Akaike 1974) of the environmental variables was applied separately for each model.

Models were then used to predict metric theoretical values in reference conditions at any site. Predictions were compared with observations and residuals (residuals = log (observations+1) – log (predictions+1)) were standardized. Assuming that most of the natural variability of the metrics was included in the models, the metric residuals were supposed to vary according to the intensity of human disturbances and independently of natural environmental variables (see Pont et al. 2006 for details).

Eleven metrics were selected regarding model quality (checking the goodness of fit and model adequacy), metric sensitivity to the different types of human disturbances and contributing non-redundant information (correlations $|r|<0.7$): trout juveniles (N%-Trout), oxyphilous species (N%-O2INTOL and S%-O2INTOL), species intolerant to habitat degradation (N%-HINTOL), species preferring to spawn in running (N%-RHPAR) or stagnant waters (S-LIPAR), tolerant (S-TOL), stenothermal (S-STTHER), omnivorous (S-OMNI) intolerant (S%-INTOL) and limnophilic species (S%-LIMNO).

Final models and predictive uncertainty

To acquire predictive uncertainty of the metrics, selected models (see supplementary material S2[2]) were implemented within a Bayesian framework (McCarthy 2007).. Bayesian inference estimates the posterior probability distribution function (PDF), $P(\theta|Y_i)$ of a set of parameters

[2] Example of Winbugs® code for the prediction of the richness metric value in quasi-undisturbed conditions, the case of *S-O2INTOL*

θ, given a set of observed data Y_i (i, the metrics) and prior probability distributions $P(\theta)$. For each metric, the PDFs of the parameters θ were estimated using the CAL dataset. A total of 100,000 iterations of the Markov chain Monte Carlo (MCMC) algorithm using Gibbs sampler were taken to approximate the posterior distributions of estimated parameters. Convergence of the MCMC chains of the model parameters was tested using the Gelman-Rubin diagnostic (Brooks and Gelman 1998) and the first 25,000 iterations were discarded as an initial burn-in period. The PDFs of the metrics' expected values were predicted at each site and were transformed as residuals, log (observations+1) – log (prediction PDF+1) and standardized as presented above. Ten thousand iterations were randomly selected to approximate the metrics' PDFs.

According to Punt and Hilborn (1997), the most commonly criticized part of a Bayesian analysis is the specification of prior distributions. Therefore non-informative priors were chosen for all the models: $\alpha_0 \sim N(1,10^6)$ and α_1 to $\alpha_n \sim N(0,10^6)$; the complete WinBUGS® codes of the log-linear models are presented in Supplementary material S2. The sensitivity to the choice of priors was tested for all parameters and Bayesian *p*-values were computed to assess the consistency between simulated and observed data (Gelman et al. 2004).

Index calculation and validation

Metric transformation – Preliminary analyses of the metrics' responses to individual human disturbances showed that the same metric could respond positively and/or negatively to different types of human disturbances. Therefore, bilateral transformations were used to obtain only negative responses of metrics to human disturbances (Oberdorff et al. 2002). Each metric (10,000 iterations) was divided by the median value of the REF sites and rescaled between zero and one to obtain values expressed as the Ecological Quality Ratio (EQR; European Union (EC) 2000).

Metric aggregation and final index – In order to maximize the sensitivity of the final index to human-induced disturbances, the six metrics showing the lowest EQR values (i.e. mean of their PDF) were retained for each site: two metrics based on abundance and four based on richness. The mean of these two groups of metrics was computed and then averaged to obtain the final index PDF. The IPR+ score was the mean of the final index distribution.

Site probability of belonging to a class status – For management purposes, the thresholds of the five ecological status classes were defined in agreement with European intercalibration rules (Willby and Birk 2010; European Communities 2011). The 10,000 final index values were derived into the five ecological status classes (high, good, moderate, poor, or bad). For each site, the proportion of values greater or lower than the moderate/good boundary was used as an estimation of the confidence in site classification.

Index responses to natural environmental variability and human disturbances - Analysis of variance procedures were used to test for the independence between the IPR+ index and environmental conditions. The IPR+ indexes' responses to human-induced disturbances were analysed using disturbance indexes (P_{Global}, P_{WQ}, P_H and P_M) and the number of strong individual disturbances ($N_{Disturb}$; i.e. the number of disturbances with high levels or "yes" in Table 1). IPR+ index uncertainties (standard deviation of the IPR+ PDF) were compared under different environmental and human disturbance conditions. All effects were tested using Kruskal-Wallis nonparametric post-hoc tests and discriminatory efficiencies were computed (DE 5% and DE 25%; Ofenbock et al. 2004).To assess the relative contribution of each metric to the IPR+ final scores, we compared the proportion of REF and highly impacted sites (P4, WQ4, H4 and M4) for which metrics were selected. Upstream (trout and grayling zones) and downstream sites (barbel and bream zones) were considered separately for this analysis.

Bayesian models were implemented using WinBUGS software (Spiegelhalter et al. 2003) and all the statistical analyses using R (version 2.13.1; R Development Core Team 2008).

Results

Selected metrics and models

The models related to the 11 selected metrics explained 20 (for N%-O2INTOL) to 67% of the total deviance (for S-TOL) (mean=45%) and Bayesian *p*-values were generally close to 0.5, showing acceptable fit to the data (see the posterior PDF of the model parameters in Supplementary material S3[3], Table S3). For all these metrics, limited differences were observed among the model parameter PDFs for different types of priors.

Metric selection within the IPR+ index calculation

[3] Parameters of the predictive models for reference conditions

For all sites, the most frequently selected metrics were S-STTHER and S%-O2INTOL for richness and N%-HINTOL and N%-RHPAR for abundance, chosen for 68, 64, 58 and 57% of the sites, respectively (Table 2). However, the most selected metrics differed for upstream (S-STTHER, S-OMNI, N%-TROUT and N%-HINTOL) and downstream sites (S-STTHER, S-LIMNO, N%-O2INTOL and N%-RHPAR). When considering only REF sites, metrics were similarly selected except for richness metrics with S-LIPAR and S-LIMNO, which were most often selected for downstream sites. Compared to REF sites, the metrics selected for heavily impacted sites were different. Nevertheless, the responses to the different pressure types were relatively comparable (Table 2). For upstream disturbed sites, N%-RHPAR, S-STTHER, S%-INTOL and S%-O2INTOL contributed generally less to the IPR+ score, whereas S-LIPAR and S%-LIMNO contributed more. For downstream disturbed sites, N%-RHPAR and S-OMNI contributed generally less to the IPR+ score, whereas N%-O2INTOL, S-STTHER and S%-INTOL contributed more.

Index responses to natural environmental variability and human disturbances

The IPR + index did not vary significantly with physiographic variables (*P*>0.05). For all the types of disturbances, IPR+ scores were significantly different for minimally and highly disturbed sites along the pressure gradients ($P \leq 0.001$; Table 3). The IPR+ score showed the strongest responses to P_{Global} and $N_{Disturb}$ with 5% discriminatory efficiencies equal to 41 and 50%, respectively (Fig. 2). Responses to single disturbances were comparable for hydrological, water quality and morphological disturbances (5% discriminatory efficiencies equal to 31, 25 and 27%, respectively). In addition, responses to P_{Global} and $N_{Disturb}$ were significant along the entire gradients. For P_{WQ} and P_M, the responses were significant in the first part of the gradient (WQ1-WQ2, M1-M2), whereas it was only significant in the middle part for P_H (H1-H3).

Index uncertainty

Uncertainty around the IPR+ score was described by the standard deviation of its PDF (SD; Fig. 3c–e). The uncertainty median was equal to 0.13 (range, 0.06–0.17) for the whole dataset. Uncertainty was on average similar for the four disturbance subsets and for upstream and downstream sites. Most of the tests were not significant for disturbance effects and environmental variable effects ($P \leq 0.05$). Nevertheless, the relationship between the IPR+ score and SD showed a bell-shaped curve (Fig. 3b), confirmed by the statistical tests with a significant decrease in uncertainty around the IPR+ score at the extremes of the disturbance gradient $N_{Disturb}$ (0–1, DE 5%=4%; 0–6, DE 5%=10%; Fig. 2b).

Confidence in site classification

Confidence in site classification as “good or better” or “moderate or worse” decreased when the IPR+ score approached the boundary (Fig. 3a). On average, the sites of the CON dataset had an 80% chance of being classified in the right side of the Good or Moderate boundary (range, 48–100%). Confidence was on average greater for highly disturbed sites (mean=87 and 89% chance for P4 sites and sites highly perturbed by more than six types of disturbance, respectively) and smaller for minimally disturbed sites (77 and 76% chance for P1 sites and one-disturbance sites; $P \leq 0.001$). In addition, in absence of disturbance, trout sites had a greater probability of being well classified (mean = 81%) than barbel and grayling sites (mean = 70 and 71%, respectively; $P \leq 0.001$). In addition, out of the 1654 CON sites, 420 (555, 845, 1147, 1407 and 1636) had more than 95% probability (90, 80, 70, 60 and 50%, respectively) of being well classified.

Discussion

The main purpose of this study was to develop a predictive multi-metric index based on fish functional guilds, relevant for French rivers and integrating uncertainties associated with reference condition predictions. The IPR+ index is based on the deviation of 11 functional metrics (four abundance metrics and seven richness metrics) from expected values in quasi-undisturbed conditions. Probability distributions of the expected metrics in absence of human pressure were predicted within a Bayesian framework from models integrating environmental variables such as temperature or the site’s geology. The IPR+ index scores and uncertainty were independent of physiographical factor variability and responded significantly to the gradients of physical and chemical disturbances. Finally, the confidence in site classification (“good and better” versus “moderate and worse”) were greater for highly perturbed sites than slightly disturbed sites.

The IPR+ index relies on metrics based on functional traits rather than taxonomic metrics for comparison of rivers and sites presenting similar ecosystem functioning but different species pools (Lamouroux et al. 2002; Hoeinghaus et al. 2007). Moreover, such metrics are generally better indicators of human-induced perturbation impacts on the biota than taxonomic metrics (Doledec et al. 2006; Marzin et al. 2012)

Our method explicitly considers the river a continuum (Vannote et al. 1980) by modelling fish assemblage structure in minimally disturbed conditions as a function of physiographical variables (Pont et al. 2006). As lowland rivers are rarely unimpaired, we advocate that moderately impacted lowland sites should be included in the calibration dataset to cover the largest gradient of environmental conditions. Obviously, this compromise decreases the index's power by removing some of the disturbance effect from the signal, but indices can be used for lowland rivers. Furthermore, this method seems satisfactory since the IPR+ scores were independent of the physiographical features, particularly those related to river size.

Almost all the models included the upstream catchment area and the climatic variables. These results are consistent with previous studies showing that hydraulic and temperature conditions are key abiotic components structuring fish communities (Blanck et al. 2007; Logez et al. 2012b). Temperature and stream power were average for the 10 years preceding the sampling, reflecting the mid-term climate condition. Such information can account for long-term climate change and its induced impacts on reference conditions (Logez and Pont 2012).

In contrast to classical bioindication tools (e.g. Pont et al. 2006), only the six most degraded metrics are selected for each site to improve the sensitivity of the IPR+ to human disturbances. Generally, intolerant metrics (S-STTHER, N%HINTOL) were most frequently selected. Metric selection does not seem to depend on the type of disturbances but on the level of disturbance and upstream-downstream position. In quasi-undisturbed conditions, the metrics selected are those naturally highly represented in fish assemblages. Intolerance metrics were selected for upstream sites while LIMNO and LIPAR were selected downstream in accordance with the longitudinal variation of fish assemblage structure (Logez et al. 2012a). By contrast, for heavily impacted sites, the metrics naturally under-represented in fish assemblages were selected: tolerant metrics for upstream sites and intolerant metrics for downstream sites. These results suggest a possible interaction between human disturbances and the environment, largely under-studied.

In agreement with previous studies (e.g. Hering et al. 2006), the IPR+ index was impacted by all the types of disturbance but more by global river degradations (i.e. a mix of different disturbances; P_{Global}, $N_{Disturb}$). More interestingly, whereas previous bioindicators usually demonstrated a lack of sensitivity to severe hydrological degradations (Marzin et al. 2012), IPR+ responses to this type of alteration were comparable to morphological and water quality degradations. Nevertheless, IPR+ responded to slight morphological and water quality degradations, although only to medium hydrological degradations, indicating that the index had different sensitivities to disturbance types.

Since the understanding of river systems is imperfect, it is essential for decision-making to recognize uncertainty and ignorance (van der Sluijs 2007). Since model adjustments and predictions relied on fish community estimation (samples) and the metrics' environmental variability was not fully explained by the model, metric values predicted for the reference condition can be uncertain. Explicitly taking into account the uncertainty due to sampling is only possible if several replicates are performed (Clarke et al. 1996). In the current study, this information was not available, but uncertainty was minimized as advocated by Angermeier et al. (2000) by considering samples that were sufficient to evaluate abundance and species richness. Nonetheless, the uncertainty associated with the metric values that would have been observed in reference conditions was successfully computed using Bayesian models and propagated through the index calculation (Bevington and Robinson 2003). The uncertainty around the IPR+ score does not seem to vary with the environment. In contrast, the bell-shaped relationship between the IPR+ standard deviation and the IPR+ scores suggests an increase in variability and uncertainty for the middle-range IPR+ scores (moderate status) and a lower uncertainty on the margins (bad and high status), perhaps partly due to the transformations necessary to acquire EQR. Unfortunately, for managers, the boundary between moderate and good classes defined the limit between degraded and good ecological status and therefore knowledge of the uncertainty around the score is essential and should be included with the index.

According to Clarke et al. (1996), the appropriate way of declaring a site's status class is by giving its probabilities of belonging to each status class. The question asked by the analysis of the confidence in site classification was: "Is the site truly above or below the good/moderate boundary?" of major importance within the scope of the WFD. Our results agree with those of Clarke et al. (1996) showing that confidence in the class increases with the distance from the class limit. The mechanical cause of this result explains why confidence in the classification could be very low while uncertainty around the score was not extremely high. Ellis (2007) discussed the rules to be used to determine if a site is truly in "good or better" or "moderate or worse" category and decided to use the Benefice-of-the-Doubt rule, defining as "good" all the sites that have less than 95% confidence of being in the "worse than moderate" class. This rule is the most indulgent and the authors felt that this decision belongs to the decision-maker. Consequently, we recommend that whatever rules are chosen to classify the sites, the confidence level should always be associated with the class or score to avoid any confusion and misinterpretation.

In conclusion, to our knowledge this is the first time that a multi-metric index has integrated the uncertainty associated with establishing reference conditions for present and future climatic conditions. In light of these results, Bayesian modelling seems an appropriate method to respond to bio-assessment issues essential for decision-making and to acquire an explicit measurement of the uncertainty around reference conditions. This new methodology is relatively generic and could be extended to other biological groups and over larger spatial extents.

Aknowledgements

This study was funded by the French National Agency for Water and Aquatic Environment (ONEMA) and the "HYNES" collaborative project of Irstea and the French Electric Company (EDF). We wish to acknowledge the contribution of the ONEMA engineers to the database compilation. We thank also Météo-France and the Hydrology Research Team, Irstea Antony for their support on climatic data.

References

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control. 19(6): 716-723.

Angermeier, P.L., Smogor, R.A., and Stauffer, J.R. 2000. Regional frameworks and candidate metrics for assessing biotic integrity in mid-atlantic highland streams. Trans. Am. Fish. Soc. **129**(4): 962-981.

Bady, P., Doledec, S., Fesl, C., Gayraud, S., Bacchi, M., and Scholl, F. 2005. Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. Freshw. Biol. **50**(1): 159-173.

Bailey, R.C., Kennedy, M.G., Dervish, M.Z., and Taylor, R.M. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. Freshw. Biol. **39**(4): 765-774.

Bevington, P.R., and Robinson, D.K. 2003. Data reduction and error analysis for the physical sciences. McGraw-Hill, New York.

Blanck, A., Tedesco, P.A., and Lamouroux, N. 2007. Relationships between life-history strategies of European freshwater fish species and their habitat preferences. Freshw. Biol. **52**(5): 843-859.

Brooks, S.P., and Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. **7**(4): 434-455.

Brooks, S.P. 2003. Bayesian computation: a statistical revolution. Phil. Trans. R. Soc. Lond. A. **361**(1813): 2681-2697. doi: 10.1098/rsta.2003.1263.

Clarke, R.T., Furse, M.T., Wright, J.F., and Moss, D. 1996. Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. J. Appl. Stat. **23**(2-3): 311-332.

Clarke, R.T., and Hering, D. 2006. Errors and uncertainty in bioassessment methods - Major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia. **566**(1): 433-439.

Doledec, S., Phillips, N., Scarsbrook, M., Riley, R.H., and Townsend, C.R. 2006. Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. J. N. Am. Benthol. Soc. **25**(1): 44-60.

Ellis, J. 2007. Combining multiple quality elements and defining spatial rules for WFD classification. Environment Agency,UK.

European Union (EC). 2000. Directive 2000/60/EC of the European Parliament and of the council establishing a framework for the community action in the field of water policy. Off. J. Eur. Commun. **L327**: 1-72.

European Communities. 2011. Implementation strategy for the Water Framework Directive (2000/60/Ec) Guidance document no. 14. Guidance document on the intercalibration process 2008-2011. European Communities.

Funtowicz, S.O., and Ravetz, J.R. 1993. Science for the post-normal age. Futures. **25**(7): 739-755.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 2004. Bayesian data analysis. Chapman and Hall, London.

Hartigan, J.A., and Wong, M.A. 1979. Algorithm AS136: A k-means clustering algorithm. Appl. Stat. **28**: 100-108.

Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K., and Verdonschot, P.F.M. 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshw. Biol. **51**(9): 1757-1785.

Hoeinghaus, D.J., Winemiller, K.O., and Birnbaum, J.S. 2007. Local and regional determinants of stream fish assemblage structure: inferences based on taxonomic vs. functional groups. J. Biogeogr. **34**(2): 324-338.

Huet, M. 1954. Biologie, profils en long et en travers des eaux courantes. Bull. fr. Piscic. **175**: 41-53.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries. **6**(6): 21-27.

Lamouroux, N., Poff, N.L., and Angermeier, P.L. 2002. Intercontinental convergence of stream fish community traits along geomorphic and hydraulic gradients. Ecology. **83**(7): 1792-1807.

Logez, M., and Pont, D. 2011. Development of metrics based on fish body size and species traits to assess European coldwater streams. Ecol. Indic. **11**: 1204-1215.

Logez, M., Bady, P., Melcher, A., and Pont, D. 2012a. A continental-scale analysis of fish assemblage functional structure in European rivers. Ecography. In press. doi: 10.1111/j.1600-0587.2012.07447.x.

Logez, M., Bady, P., and Pont, D. 2012b. Modelling the habitat requirement of riverine fish species at the European scale: sensitivity to temperature and precipitation and associated uncertainty. Ecol. Freshw. Fish. **21**(2): 266-282.

Logez, M., and Pont, D. 2012. Global warming and potential shift in reference conditions: the case of functional fish-based metrics. Hydrobiologia. In press.

Marzin, A., Archaimbault, V., Belliard, J., Chauvin, C., Delmas, F., and Pont, D. 2012. Ecological assessment of running waters: do macrophytes, macroinvertebrates, diatoms and fish show similar responses to human pressures? Ecol. Indic. **23**: 56-65.

McCarthy, M.A. 2007. Bayesian methods for ecology. Cambridge University Press, Cambridge.

McCormick, F.H., Hughes, R.M., Kaufmann, P.R., Peck, D.V., Stoddard, J.L., and Herlihy, A.T. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Trans. Am. Fish. Soc. **130**(5): 857-877.

McCullagh, P., and Nelder, J.A. 1989. Generalized linear models. Chapman and Hall, London.

Noble, R.A.A., Cowx, I.G., Goffaux, D., and Kestemont, P. 2007. Assessing the health of European rivers using functional ecological guilds of fish communities: standardising species classification and approaches to metric selection. Fish. Manag. Ecol. **14**(6): 381-392.

Oberdorff, T., Pont, D., Hugueny, B., and Porcher, J.P. 2002. Development and validation of a fish-based index for the assessment of 'river health' in France. Freshw. Biol. **47**(9): 1720-1734.

Ofenbock, T., Moog, O., Gerritsen, J., and Barbour, M. 2004. A stressor specific multimetric approach for monitoring running waters in Austria using benthic macro-invertebrates. Hydrobiologia. **516**(1-3): 251-268.

Ostermiller, J.D., and Hawkins, C.P. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. J. N. Am. Benthol. Soc. **23**(2): 363-382.

Pont, D., Hugueny, B., and Oberdorff, T. 2005. Modelling habitat requirement of European fishes: do species have similar responses to local and regional environmental constraints? Can. J. Fish. Aquat. Sci. **62**(1): 163-173.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., and Schmutz, S. 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. J. Appl. Ecol. **43**(1): 70-80.

Pont, D., Hugueny, B., and Rogers, C. 2007. Development of a fish-based index for the assessment of river health in Europe: the European Fish Index. Fish. Manag. Ecol. **14**(6): 427-439.

Poulet, N., Beaulaton, L., and Dembski, S. 2011. Time trends in fish populations in metropolitan France: insights from national monitoring data. J. Fish Biol. **79**(6): 1436-1452. doi: 10.1111/j.1095-8649.2011.03084.x.

Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. **7**(1): 35-63.

R Development Core Team. 2008. R: A language and environment for statistical computing Vienna, Austria.

Spiegelhalter, D.T.A., Best, N., and Lunn, D. 2003. WinBUGS user manual. Version 1.4. Available at http://www.mrc-bsu.cam.ac.uk/bugs.

Staniszewski, R., Szoszkiewicz, K., Zbierska, J., Lesny, J., Jusik, S., and Clarke, R.T. 2006. Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification. Hydrobiologia. **566**(1): 235-246.

Tenenhaus, M., and Young, F.W. 1985. An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data. Psychometrika. **50**: 91-119.

van der Sluijs, J.P. 2007. Uncertainty and precaution in environmental management: insights from the UPEM conference. Environ. Modell. Softw. **22**(5): 590-598.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., and Cushing, C.E. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. **37**(1): 130-137.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M. 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. Int. J. Climatol. **30** (11): 1627-1644. doi: 10.1002/joc.2003.

Willby, N., and Birk, S. 2010. IC Guidance Annex V: definition of comparability criteria for setting class boundaries. WG ECOSTAT.

Table 1. Description of human-induced disturbances at the 1654 French sites and site selection criteria. For degradation classes, increasing numbers indicate increasing disturbances (e.g. WQ1, WQ2, WQ3, WQ4). † Variables used in the global disturbance indices (P_{Global} and $N_{Disturb}$), * Variables used in the hydrological disturbance index (P_H), ‡ Variables used in the morphological disturbance index (P_M), § Variables used in the water quality disturbance index (P_{WQ})

Variables	Categories of the variables (number of sites)	REF	CAL (Trout zone)	CAL (Grayling zone)	CAL (Barbel zone)	CAL (Bream zone)
Downstream barrier †	No (896) / Partial (337) / Yes (370) / NA (51)	No-Partial	No-Partial	No-Partial	No-Partial	No - Partial
Hydrological regime modified * †	No (803) / Slight (408) / Moderate (276) / High (167)	No-Slight	No	No-Slight	No-Slight	No - Slight
Hydropeaking * †	No (1355) / Slight (159) / Moderate (68) / High (72)	No-Slight	No	No-Slight	No-Slight	No - Slight
By-pass channel * †	No (1482) / Slight (104) / Moderate (54) / High (14)	No-Slight	No	No-Slight	No-Slight	No - Slight
Water abstraction * †	No (808) / Slight (580) / Moderate (131) / High (135)	No-Slight	No	No-Slight	No-Slight	No - Slight
Presence of an impoundment *†	No (1290) / Slight (172) / Moderate (44) / High (148)	No-Slight	No	No-Slight	No-Slight	No - Slight
Artificial embankment ‡	No (1174) / Slight (278) / Moderate (112) / High (90)					
Riparian vegetation modified ‡	No (850) / Slight (437) / Moderate (207) / High (160)					
Sedimentation ‡	No (667) / Slight (547) / Moderate (266) / High (174)					
Channel form modified † ‡	No (1132) / Partial (303) / Yes (219)	No- Partial	No	No- Partial	No- Partial	No - Partial
Cross-section modified † ‡	No (1065) / Partial (318) / Yes (271)	No- Partial	No	No- Partial	No- Partial	No - Partial
Channel incision or aggradation ‡	No (1063) / Partial (412) / Yes (179)					
Water temperature modified	No (1259) / Cooling (56) / Warming (299) / NA (40)					
Toxic pollution	No (542) / Slight (505) / Moderate (132) / High (117) / NA (358)					
Organic pollution †§	No (404) / Slight (481) / Moderate (152) / High (86) / NA (531)	No-Slight	No-Slight	No-Slight	No-Slight	No - Slight
Nutrient pollution §	No (277) / Slight (457) / Moderate (258) / High (72) / NA (590)					
Eutrophication †§	No (710) / Slight (394) / Moderate (169) / High (68) / NA (313)	No-Slight	No	No-Slight	No-Slight	No - Moderate
Acid pollution	No (1120) / Yes (35) / NA (499)					
Organic pollution SEQ Class †§	1(328) / 2 (670) / 3 (327) / 4 (141) / 5 (87) / NA (101)	1-2	1-2	1-2	1-2-3	1 - 2
Upstream lake †	No (1183) /Partial (256) / Yes (153) / NA (74)	No - Partial	No	No - Partial	No - Partial	No - Partial
Navigation †	No (1515) /Partial (31) / Yes (39) / NA (74)	No - Partial				
Water quality degradation index (P_{WQ})	WQ1 (160) / WQ2 (77) / WQ3 (89) / WQ4 (128) / NA (1200)					
Morphological degradation index (P_M)	M1 (425) / M2 (167) / M3 (109) / M4 (149) / NA (804)					
Hydrological degradation index (P_H)	H1 (316) / H2 (206) / H3 (124) / H4 (140) / NA (868)					
Number of disturbance types ($N_{Disturb}$)	0 (400) / 1 (299) / 2 (228) / 3 (167) / 4 (143) / 5 (113) / > 6 (262)					
Global disturbance index (P_{Global})	P1 (359) / P2 (386) / P3 (328) / P4 (357) / NA (224)					

Table 2. Frequency of the metrics selected within the IPR+ index calculation. For each site, the four most impacted richness metrics (S and S%) and the two most impacted abundance metrics (N%) are selected to calculate the IPR+ index. The percentage of sites where a given metric is selected is compared to the reference condition sites (REF dataset) and sites highly impacted by different types of human disturbances (P4 = global disturbance, WQ4 = water quality degradation, H4 = hydrological degradation, M4 = morphological degradation). *P*-value of the proportion comparisons (*** ≤0.001, ** ≤0.01 and * ≤0.05). > for increasing selection, < for decreasing selection.

	Nsites	N%-TROUT	N%-O2INTOL	N%-HINTOL	N%-RHPAR	S-TOL	S-STTHER	S-LIPAR	S-OMNI	S%-INTOL	S%-O2INTOL	S%-LIMNO
CON dataset	1654	-	48%	58%	57%	59%	68%	47%	58%	54%	64%	51%
REF dataset	266	-	35%	56%	62%	62%	73%	32%	70%	64%	73%	27%
Upstream sites (trout and grayling zones, Huet 1954)												
CON sites	857	72%	23%	53%	52%	69%	71%	31%	71%	60%	64%	35%
REF sites	188	66%	23%	53%	58%	65%	81%	18%	74%	75%	76%	10%
P4 sites	71	ns	> ***	ns	< ***	ns	< ***	> ***	< ***	< ***	< *	> ***
WQ4 sites	56	ns	ns	ns	< *	ns	ns	> ***	ns	< ***	< ***	> ***
H4 sites	62	ns	ns	ns	< ***	> *	< **	> ***	ns	< ***	< ***	> ***
M4 sites	54	ns	ns	ns	< **	> *	< ***	> ***	ns	< **	< ***	> ***
Downstream sites (barbel and bream zones, Huet 1954)												
CON sites	797	-	75%	63%	63%	49%	64%	63%	44%	47%	63%	69%
REF sites	78	-	63%	64%	73%	54%	53%	67%	58%	36%	65%	68%
P4 sites	286	-	> ***	ns	< ***	ns	> ***	ns	< ***	> ***	ns	< *
WQ4 sites	72	-	> *	ns	< **	ns	ns	ns	ns	ns	ns	ns
H4 sites	78	-	> *	ns	< ***	ns	> **	ns	< **	> *	ns	ns
M4 sites	95	-	> ***	ns	< ***	ns	> *	ns	< *	> *	ns	ns

Table 3. IPR+ index responses to human disturbances (score and uncertainty). DE 5% (DE 25%), discriminatory efficiency: the percentage of sites impacted (e.g. P4, PC4, M4 or H4) with IPR+ score less than the 5% (25%) extreme percentile of the minimally disturbed sites (number of disturbances = 0, P1, PC1, M1 and H1). *p*-value: *** ≤0.001, ** ≤0.01 and * ≤0.05 and *ns* for nonsignificant test.

	IPR+ score (mean)			IPR+ uncertainty (SD)		
	DE 5%	DE 25%	*p*-value	DE 5%	D 25%	*p*-value
Number of human disturbances ($N_{Disturb}$)						
0-1	8%	40%	***	4%	18%	**
0-2	19%	50%	***			*ns*
0-3	12%	53%	***			*ns*
0-4	24%	63%	***			*ns*
0-5	30%	68%	***			*ns*
0- > 6	**50%**	**76%**	*******	**10%**	**35%**	*******
Global disturbance index (P_{Global})						
P1-P2	10%	44%	***			*ns*
P1-P3	20%	59%	***			*ns*
P1-P4	**41%**	**76%**	*******	**7%**	**27%**	******
Water quality degradation index (P_{WQ})						
WQ1-WQ2	10%	40%	*	1%	13%	**
WQ1-WQ3	26%	55%	***			*ns*
WQ1-WQ4	**25%**	**54%**	*******	**2%**	**13%**	*******
Hydrological degradation index (P_H)						
H1-H2			*ns*			*ns*
H1-H3	22%	51%	***			*ns*
H1-H4	**31%**	**61%**	*******			***ns***
Morphological degradation index (P_M)						
M1-M2	14%	40%	***			*ns*
M1-M3	18%	54%	***			*ns*
M1-M4	**28%**	**61%**	*******			***ns***

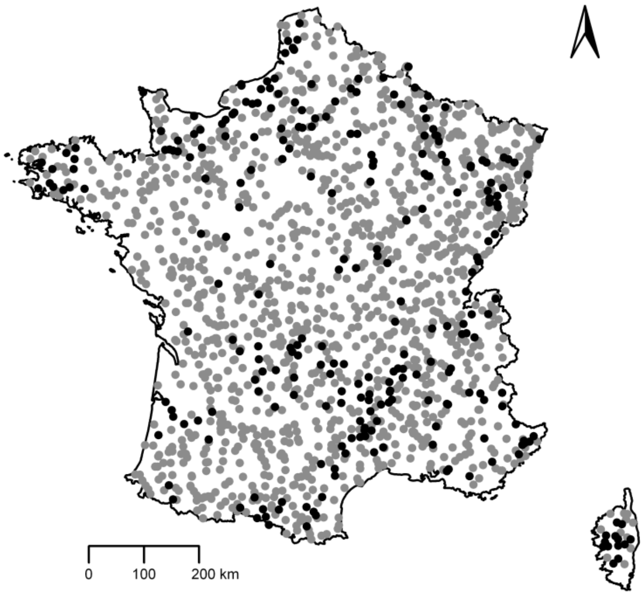


Fig. 1. Location of the 1654 French sites. Dark grey, calibration sites (CAL dataset); light grey, the other sites of the construction dataset (CON dataset).

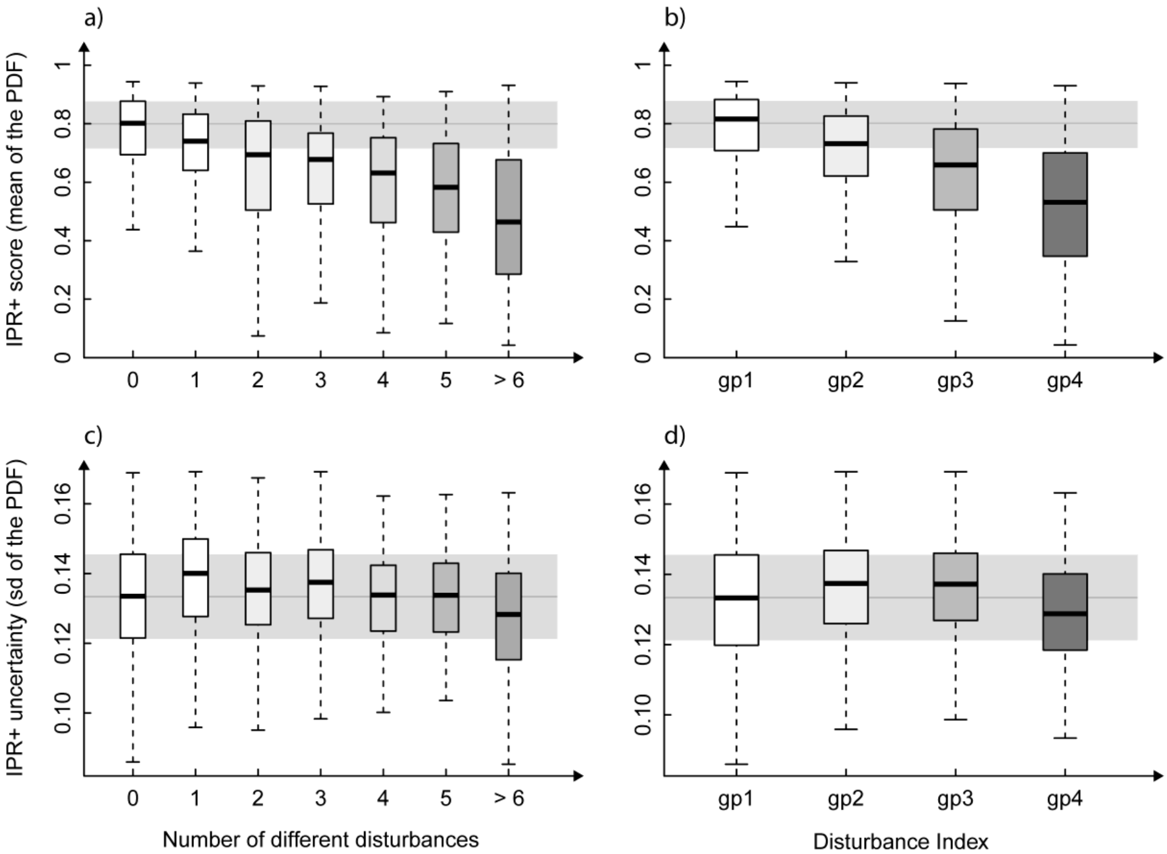


Fig. 2. IPR+ (probability distribution function, PDF, 10,000 iterations) responses to the global disturbance indexes. (a) and (b) IPR+ score (mean of the PDF) and (c) and (d) IPR+ uncertainty (SD of the PDF). (a) and (c) Responses to the number of different strong human disturbances at a single site (from zero to six and more types). (b) and (d) responses to the global disturbance index (P_{Global}; from P1 = slight modifications to P4 = high modifications).

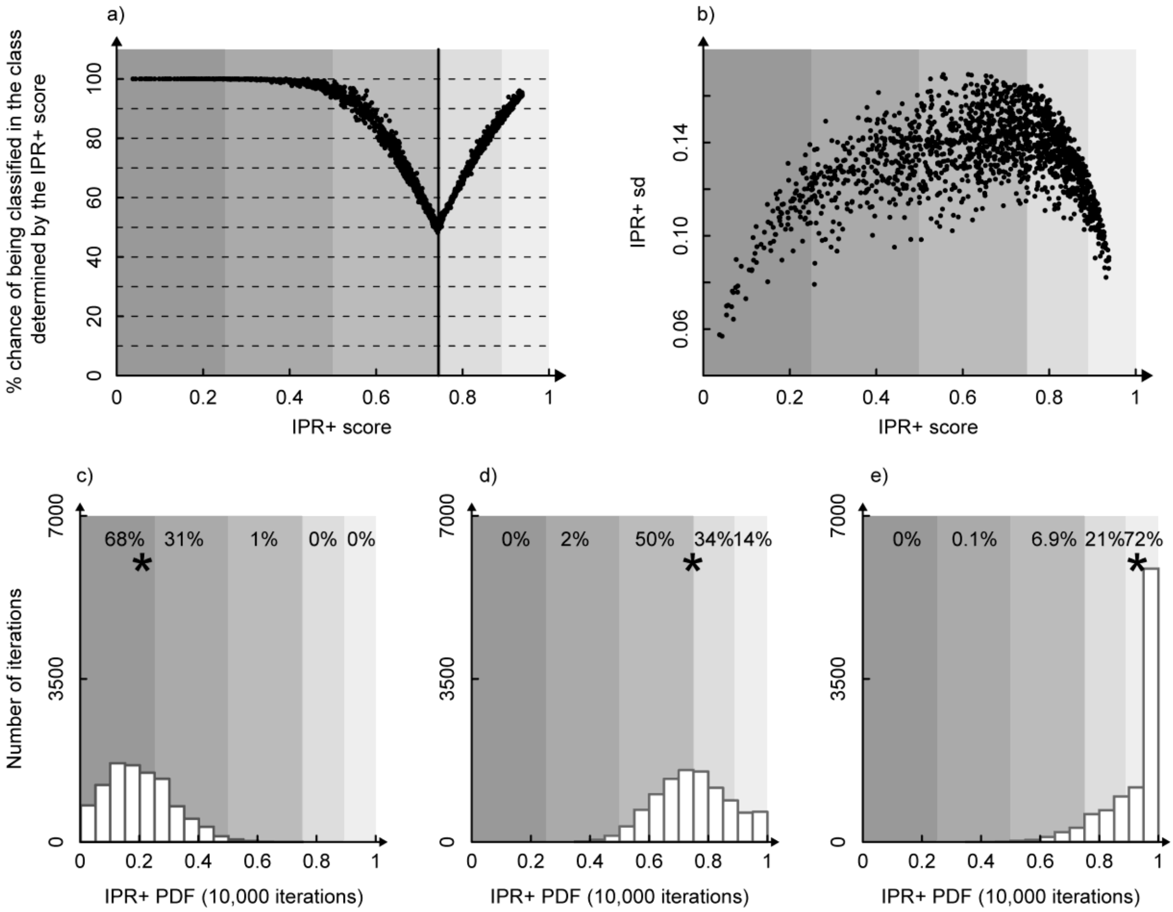


Fig. 3. IPR+ probability distribution function (PDF; 10,000 iterations) and confidence in classification "better than good" versus "worse than moderate". In different grey tones, from left to right the five ecological status classes: (bad, poor, moderate, good, high). (a) Percentage of chances (iterations) to be classified in the right side of the good/moderate boundary as a function of the IPR+ score. (b) IPR+ uncertainty (standard deviation of the PDF) in function of the IPR+ score (mean of the PDF). (c)–(e) Examples of IPR+ PDF for three river stations: (c) the river Slack at Ambleteuse, (d) the river Oise at Macquigny and (e) the river Sierre at Montcel. Percentages correspond to the probability of the station being in each of the five status classes. Black stars indicate IPR+ scores (mean of the PDF).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. French fish biological and ecological functional traits

Table 1. Description of the 12 biological and ecological traits

Appendix S2. Example of Winbugs® code for the prediction of the richness metric value in quasi-undisturbed conditions, the case of *S-O2INTOL*

Appendix S3. Parameters of the predictive models for reference conditions

Table 1. Selected metrics and predictive models

Appendix S1. French fish biological and ecological functional traits

Table 1: Description of the 12 biological and ecological traits

Traits	Categories
Tolerance to water quality degradation (WQ)	Intolerant (INTOL)
	Intermediate (IM)
	Tolerant (TOL)
Tolerance to oxygen (O_2)	Intolerant (O2INTOL): species requiring more than 6 mg of oxygen per litre
	Intermediate (O2IM): species relatively tolerant to low oxygen concentration
	Tolerant (O2TOL): species able to live in water with less than 3 $mg.L^{-1}$.
Temperature tolerance (TEMP)	Eurythermal (EUTHER): species able to withstand a wide range of temperature
	Stenothermal (STTHER): species able to withstand a narrow range of temperature
Tolerance to habitat degradation (HAB)	Intolerant (HINTOL)
	Intermediate (HIM)
	Tolerant (HTOL)
Affinity to flow velocity (VEL)	Limnophilic (LIMNO): species preferring to live in slow-flowing to stagnant conditions
	Rheophilic (RH): species preferring to live in high-flow conditions
	Eurytopic (EURY): species with a wide tolerance to flow conditions
Feeding habitat (FHAB)	Benthic (B): species preferring to live near the bottom from where they feed
	Water column (WC): species that live and feed in the water column
Adult trophic guild (TROPH)	Detritivorous (DETR): adult diet composed of a high proportion of detritus
	Herbivorous (HERB): adult diet is composed of at least 75% plant material
	Insectivorous (INSV): adult diet is composed of at least 75% insect individuals
	Omnivorous (OMNI): adult diet is composed of more than 25% plant material and more than 25% animal material
	Piscivorous (PISC): adult diet composed of more than 75% fish
	Planktivorous (PLAN): adult diet is composed of more than 75% phytoplankton or zooplankton
Migration behaviour (MIG)	Anadromous (LMA): species living as older juveniles and sub-adults in the sea and migrating up rivers to spawn at maturity
	Catadromous (LMC): species with early life stage living in fresh water and migrating down rivers to spawn in the sea at maturity
	Resident (RESID): species moving over small areas within particular river segment
	Potamodromous (POTAD): species migrating within the inland waters of a river
Reproduction (REPRO)	Lithophilic (LITHO): species spawning exclusively on gravel, rocks, stones, rubbles or pebbles and with photophobic hatchlings
	Ostracophilic (OSTRA): species spawning in bivalve molluscs
	Phyto-lithophilic (PHLI): species depositing their eggs in clear water habitats on submerged plants or on other submerged items such as logs, gravel and rocks and their larvae are photophobic
	Phytophilic (PHYT): species depositing their eggs in clear water habitats on submerged plants
	Viviparous (VIVI) viviparous species
Spawning habitat (RHAB)	(LIPAR) species preferring to spawn in stagnant water
	(RHPAR) species preferring to spawn in running waters
	(EUPAR) species without clear spawning preferences
Reproductive behaviour (R)	Single (SIN): species with a single spawning event during the reproductive season
	Fractional (FR): species which either spawn repeatedly in a season or with different components of their populations spawning at different times
	Protracted (PRO): species spawning over a long period during the reproductive season
Parental care (PC)	(PROT) species presenting egg or larva life stages with protection
	(NOP) species with no protection for early life stages

Appendix S2. Example of Winbugs® code for the prediction of the richness metric value in quasi-undisturbed conditions, the case of *S-O2INTOL*.

```
# log-linear models for richness metrics (number of species)

model {

        # PRIOR (non informative)
        alpha ~ dnorm(1,1.0E-6) # intercept
        alpha_lCA ~ dnorm(0.0,1.0E-6) # parameter lCA
        alpha_lCA2 ~ dnorm(0.0,1.0E-6)
        alpha_lPOW   ~ dnorm(0.0,1.0E-6)
        alpha_lPOW2   ~ dnorm(0.0,1.0E-6)
        alpha_AT  ~ dnorm(0.0,1.0E-6)
        alpha_AT2  ~ dnorm(0.0,1.0E-6)
        alpha_ATA  ~ dnorm(0.0,1.0E-6)


        # LIKELIHOOD
        for(i in 1 : C) {  # for each sites

        # log linear model
        log(mu[i]) <- alpha + alpha_lPOW * lPOW[i] + alpha_lPOW2 *(lPOW[i]* lPOW[i])
                + alpha_lCA * lCA[i] + alpha_lCA2 * (lCA [i]* lCA [i])
                + alpha_AT * AT[i] + alpha_AT2 * (AT[i]* AT[i])
                + alpha_ATA * ATA[i]
                + lS_TOT[i] # offset with the total number of species for relative richness S%

        S_O2INTOL[i] ~ dpois(mu[i])
        }


        # PREDICTIONS FOR NEW SITES
         for (n in 1:N) { # for each sites
         log(mu_New [n]) <- alpha
        + alpha_lPOW * lPOW_New[n] + alpha_lPOW2 *(lPOW_New[n]* lPOW_New[n])
                + alpha_lCA * lCA_New[n] + alpha_lCA2 * (lCA_New[n]* lCA_New[n])
                + alpha_AT * AT_New[n] + alpha_AT2 * (AT_New[n]* AT_New[n])
                + alpha_ATA * ATA_New[n]
                + lS_TOT_New[n]

        S_O2INTOL_New[n] ~ dpois(mu_New[n])
        }


}
```

Appendix S3. Predictive models for reference conditions

Table 1: Selected metrics and predictive models

Metrics	Bayesian parameter estimations (10,000 iterations)																							
	Intercept		**lPOW**		**lPOW^2**		**lCA**		**lCA^2**		**AT**		**AT^2**		**ATA**		**ATA^2**		**GT_S**		**H_PM**		**H_NG**	
	Mean	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*	***Mean***	*Sd*
N%-TROUT	***-9.301***	*0.895*	***0.428***	*0.079*			***1.067***	*0.280*	***-0.217***	*0.035*					***0.155***	*0.054*								
N%-O2INTOL	***-1.479***	*0.475*	***0.158***	*0.037*			***-0.032***	*0.107*	***-0.027***	*0.010*														
N%-HINTOL	***1.676***	*0.954*	***0.213***	*0.099*	***0.002***	*0.005*	***-0.433***	*0.136*	***0.006***	*0.012*	***-0.063***	*0.037*			***-0.366***	*0.129*	***0.012***	*0.005*	***-0.215***	*0.112*	***0.252***	*0.115*	***0.078***	*0.190*
N%-RHPAR	***-1.095***	*0.295*	***0.066***	*0.017*			***0.054***	*0.048*	***-0.017***	*0.005*					***0.024***	*0.011*								
S-TOL	***4.992***	*1.631*	***-1.559***	*0.348*	***0.055***	*0.016*	***1.348***	*0.168*	***-0.063***	*0.014*	***0.062***	*0.031*							***-0.247***	*0.091*	***0.276***	*0.113*	***-0.003***	*0.212*
S-STTHER	***3.819***	*1.534*					***0.363***	*0.117*	***-0.041***	*0.012*	***0.641***	*0.175*	***-0.033***	*0.008*	***-0.899***	*0.137*	***0.029***	*0.005*			***0.145***	*0.109*	***-0.324***	*0.193*
S-LIPAR	***2.531***	*2.115*	***-1.958***	*0.251*	***0.065***	*0.011*	***1.430***	*0.203*	***-0.054***	*0.015*					***0.355***	*0.203*	***-0.005***	*0.007*	***-0.477***	*0.150*	***0.616***	*0.186*	***0.251***	*0.351*
S-OMNI	***-8.079***	*1.892*	***-1.139***	*0.225*	***0.040***	*0.010*	***1.472***	*0.146*	***-0.078***	*0.011*	***0.904***	*0.242*	***-0.037***	*0.010*	***0.562***	*0.059*	***-0.014***	*0.002*						
S%-INTOL	***2.206***	*1.177*	***0.175***	*0.037*			***-0.248***	*0.029*			***-0.045***	*0.029*			***-0.461***	*0.134*	***0.016***	*0.005*						
S%-O2INTOL	***-0.012***	*0.975*	***-0.325***	*0.155*	***0.025***	*0.007*	***0.094***	*0.105*	***-0.046***	*0.011*	***0.279***	*0.095*	***-0.016***	*0.005*	***-0.040***	*0.020*								
S%-LIMNO	***-5.180***	*1.180*	***-1.751***	*0.219*	***0.057***	*0.010*	***1.550***	*0.185*	***-0.067***	*0.013*					***1.053***	*0.203*	***-0.030***	*0.007*						

P4

Temporal variability and uncertainty associated to river ecological assessment: the case of fish communities in French rivers.

Marzin, A., M. Logez & D. Pont.

In preparation for submission to *Ecological Indicators*

Abstract

The context of this study was the validation of a bio-indicator developed for the evaluation of French river ecological health using fish communities (IPR+). In order to acknowledge the variability of the IPR+ index unrelated with human-degradation and give some recommendations for potential users, the inter-annual variability of the IPR+ index and underlying metrics was quantified and analysed while accounting for their predictive uncertainty. IPR+ index, metrics and associated uncertainty were computed for ten-year time series (1998-2007) at 183 French sites. Reliability of the evaluation and inter-annual variations were quantified and analysed using statistical methods.

IPR+ scores of the 183 sites were mostly consistent through time (Pearson correlations, 0.74-0.88). Uncertainty of the metric values and IPR+ score was generally much larger (SD_U, 0.14-0.35) than the inter-annual variability (SD_T, 0.6-0.14) explaining the weak numbers of sites showing significant inter-annual variations of IPR+ index and metrics. Finally, IPR+ index showed comparable inter-annual variations along the longitudinal gradient and was significantly less stable at degraded sites than at more pristine sites. Accordingly, in addition to weak values of the IPR+ index, the lack of the index inter-annual stability might be an evidence of impacts of human disturbances on river fish communities.

Finally, our results confirm the advantages of multi-metrics index based on ecological and biological functional metrics for the assessment of river ecological health and endorse the need for sampling reference sites through time to account for reference conditions shift and global changes for future bio-assessments.

1. Introduction

As the world population grows over time, industrialization and urbanization follows rapid evolutions, drinking water supply, sustainable energy, industrial, and agricultural interests have increased and competitions between river users are exacerbated (Huang and Chang, 2003; Wang *et al.*, 2003). In this context, water managers need efficient and complete tools to assess the impacts of human activities on aquatic ecosystems. Biological indicators are widely used to evaluate the water bodies' ecological quality (Palmer *et al.*, 2005). The objective of bio-assessment is to interpret degradations of ecosystems related to human disturbances by observing the resident biological communities. Bio-indicators are often based on the Reference Condition Approach (RCA; Bailey *et al.,* 1998) for which the deviation between the current structure of aquatic assemblages and the structure expected in condition quasi-undisturbed by human activity (Stoddard *et al.,* 2006) represents the degree of degradation of a given site. The characteristics of the expected communities could be predicted trough statistical models based on environmental features (e.g. Oberdorff *et al.*, 2002; Pont *et al.*, 2006).

As pointed out by Clarke and Hering (2006), there are multiple sources of variability in water body health assessment such as temporal variation of communities, spatial variation within water bodies and uncertainty associated with sampling processing methods and modelling errors. They all represent possible risk for decision makers to misclassify sites and therefore need to be acknowledged to provide operational decision tools.

Few studies have attended to measure the uncertainty linked to bio-indication tools assessment and most of them focused on the uncertainty linked to the sampling design and effort (e.g. Bady *et al.*, 2005; Ostermiller and Hawkins, 2004) and did not recognized the particular effect of temporal variability. Other authors implicitly recognized the effect of temporal variability and considered it in the process of metric selection (e.g. Hughes *et al.*, 2004). A metric is inappropriate if its "temporal variability does not allow discrimination between anthropogenic influences and natural variability" (Hering *et al.*, 2006). In a recent study, Marzin *et al.* (*submitted*) developed a new methodology to quantify the uncertainty of metric values predicted in reference conditions. This methodology enabled to assess the uncertainty of a fish-based index designed for French streams: IPR+. The predictive uncertainty of IPR+ index was invariant with physiographical features and was smaller for highly and not impacted conditions than for middle range of human-induced degradations.

One of the assumptions of the bio-assessment approaches is that in absence of anthropogenic disturbances, biological communities change little between years. However, responses of aquatic communities to environment are likely to change over time through biological processes, such as competition, predation or recruitment (Bunn and Davies, 2000; Ricklefs and Schluter, 1993). Therefore, assemblage response to human-induced changes may be confused with natural dynamic of aquatic communities and populations.

In addition, as the probability of detecting a species is known to increase with the accumulation of sampling effort (Gotelli and Colwell, 2001), temporal variability of bio-indicators may also encompass between year variation in sampling effort.

Nevertheless, the relative contributions of these various sources of inter-annual variation (e.g. random sampling variation, environmental stochasticity) are not well understood (Pyron *et al.*, 2008). Since evaluations generally rely on sites sparsely sampled along time, the recognition and a better understanding and of annual temporal variation is necessary to examine the likelihood of detecting human-induced perturbations (Growns *et al.*, 2006).

Whereas temporal variability of river biological communities related to the environment have been widely documented (e.g. Franquet *et al.*, 1995; Bady *et al.*, 2004; Growns *et al.*, 2006; Schaeffer *et al.*, 2012), rare studies attempted to quantify the temporal variability of bio-indicator indices and they often focused on within year variability rather than inter-annual variability (Carlisle and Clements, 1999; Fore et *al.*, 1994; Linke *et al.*, 1999; Mazor *et al.*, 2009; Collier, 2008; Pyron *et al.*, 2008). It seems that the sensitivity of multimetric indices to inter-annual variability depends on the inherent characteristics of community (e.g. diversity; Ross *et al.*, 1985), stream size (Fore *et al.*, 1994), frequency of disturbance regimes (Franssen *et al.*, 2011; Schaeffer *et al.*, 2012) and impairment level (Fore *et al.*, 1994).

To follow the WFD requirements and as a further development of the French fish index (IPR; Oberdorff *et al.*, 2001; 2002), the new river French fish index (IPR+) was built and associated with a measure of the predictive uncertainty (Marzin *et al.*, *submitted*). In this study, we aimed to quantify and understand the temporal variability of the IPR+, to provide a complete decision tool for river managers. Two main questions were discussed: in stable environmental and human disturbance conditions:

(i) How variable are the IPR+ and the underlying metrics at individual sites through time, i.e. what is the relative importance of inter-annual variability versus uncertainty?

(ii) What is the influence of disturbance intensity and longitudinal gradient (upstream-downstream) on IPR+ temporal variability?

Uncertainty and temporal variability of the IPR+ index and the associated metrics were quantified for 183 French sites using Bayesian statistical models (Marzin *et al.*, *submitted*) and temporal variability indices.

2. Material and methods

2.1. Time series data

In this study, 183 French sites for which no detectable change in human disturbance occurred between 1998 and 2007 were selected (TEMP data set; Fig. 1; Table 1). Each year of the period 1998-2007, fish communities of these sites were sampled one time by electrofishing during the low-flow period (national monitoring programs; Poulet *et al.*, 2011). To homogenize the sampling effort, only fish collected during the first pass were considered. Rare species were removed from samples (N < 2 for sites with less than 200 fish caught, N < 3 for other sites). To limit sampling error, only sites with more than 50, 100 and 200 fish caught and a fished area greater than 100 m² were selected for the trout, grayling-barbel and bream zones respectively (Huet, 1954).
Environmental factors known to influence fish assemblages (Pont *et al.*, 2005; Logez *et al.*, 2012) were described. Streams ranged from small to large size (width from 2.5 to 300 m, median = 8.6 m; distance to the source from 2 to 893 km, median = 28 km), were part of small to relatively large catchments (from 4 to 110,248 km^2, median = 224 km^2) and were situated from 5 to 1800 m above sea level (median = 162 m). Climatic data, the mean annual air temperature (AT, from 6 to 15 °C, median = 11 °C), the annual air temperature amplitude (ATA, from 10 to 19 °C, median = 16 °C) and annual precipitation in the catchment (PREC, from 626 to 2,107 mm/year, median = 991 mm/year) were obtained from SAFRAN models provided by Météo-France, the French meteorology institute (Vidal *et al.*, 2009). Geology (siliceous = 110 sites, calcareous = 73 sites) and hydrological regime (pluvial strong, pluvial moderate and nival-glacial, N = 54, 117 and 12 sites, respectively) were also described for each site. These variables were assumed to be not or weakly influenced by anthropogenic activities.

Human-induced alterations of rivers were assessed considering stream morphology (e.g. instream habitat, channel form modifications), hydrology (e.g. water abstraction, hydropeaking), water quality and connectivity (presence of barrier downstream) (Table 1). The first axis of a multivariate analysis summarizing human disturbances variables (multiple correspondence analysis; Tenenhaus and Young, 1985) was used as a global human disturbance gradient (P_G). P_G account for 11.3% of the total inertia and was related to an increase in disturbances. Five classes (Gp) of disturbance level ranging from one (not or slightly impacted sites) to five (heavily impacted sites) were defined by K-means algorithm (Hartigan and Wong, 1979; Marzin *et al.*, 2012) (N=45, 35, 33, 42, 28 for Gp1 to Gp5).

2.2. IPR+ metrics and index computation

Ten functional metrics and one metric reflecting the size structure of brown trout (*Salmo trutta*) populations are involved in the IPR+ index computation (metrics expressed in relative abundance, %N, richness S and relative richness %S): trout juveniles (N%-Trout), oxyphilous species (N%-O2INTOL and S%-O2INTOL), species intolerant to habitat degradation (N%-HINTOL), species preferring to spawn in running (N%-RHPAR) or stagnant waters (S-LIPAR), tolerant (S-TOL), stenothermal (S-STTHER), omnivorous (S-OMNI) intolerant (S%-INTOL) and limnophilic species (S%-LIMNO). N%-Trout was only applied for upstream sites (N= 99). IPR+ is a predictive multi-metric index based on the deviation between metric expected values in minimally disturbed condition (MCA; Stoddard *et al.*, 2006) and metric observed values. Expected metric values in MCA are predicted from statistical models integrating environmental variables such as the geology or the temperatures and catchment precipitations of the ten year preceding the sampling. Consequently, climatic variations were taken into account by the models and allowed to consider the changes in long- and mid-term reference conditions. IPR+ index, metrics and uncertainty associated were calculated for the 183 time series of the TEMP dataset following Marzin *et al.* (*submitted*). Hereafter, metric names refer to the transformed metrics as integrated in the final index computation (deviation from reference and ecological quality ratio, see Marzin *et al.*, *submitted*). In the IPR+ computation process, probability distribution functions (PDF; 10,000 iterations) of the index and metrics were estimated using Bayesian modelling. Index and metrics values were defined as the means of the PDFs and the uncertainties of the index and metrics were described by the standard deviations (SD_U) of the PDFs.

2.3. Data analysis

To assess the global consistency of the IPR+ evaluation through time, all pairwise Pearson correlations of the IPR+ scores and metrics values between years were calculated. To test the significance of inter-annual variations of the IPR+ scores, it was indispensable to identify whether the ten-year variability of the scores was substantial with regard to the uncertainty of individual score values. For each site, the number of the ten IPR+ scores (1998-2007) that were significantly different from the 10-year median score was computed (no overlap of the PDF 2.5 and 97.5% percentiles with the median value). Also, for each site, year-wise comparisons of the IPR+ score distributions was performed (for a pair of years, no significant overlap between the two PDFs, 2.5 and 97.5% percentiles). Same analyses were conducted for the eleven metrics.

In addition, as the IPR+ index score depends on underlying metric values and determines the ecological status class of a site, it was fundamental to analyse and quantify their respective temporal variations. The temporal variations of scores and metric values were quantified by the means of their inter-annual standard deviation (SD_T) at each site. Finally, IPR+ and metrics SD_T were compared along the gradient of human disturbance (P_G gradient; Gp1-2, Gp3, Gp4-5) and along the upstream-downstream gradient (i.e. Trout-Grayling and Barbel-Bream zones; Huet, 1954), using one-way analysis of variance (ANOVA). N%-Trout was not calculated for downstream sites and consequently only the P_G effect was tested.

Bayesian models were implemented using WinBUGS software (Spiegelhalter *et al.,* 2003) and all the statistical analyses using R (version 2.13.1; R Development Core Team, 2011).

3. Results

3.1. IPR+ index and metrics inter-annual variability vs. predictive uncertainty

Evaluations of the ecological conditions of the183 sites, using IPR+ index, were consistent over the 10 years (Table 2; e.g. Fig. 2a, b). IPR+ scores were well correlated among the different years with inter-annual correlations (*r*) in average equal to 0.82 and ranging from 0.74 to 0.88.

Uncertainty of the index ($SD_{U;}$ mean=0.14) was in average twice larger than inter-annual variability (SD_T; mean=0.07) and both were very variable among sites (0.04 - 0.17 and 0.01 -

0.17, respectively) (Table 2). Nevertheless, SD_U varied little among years, the average standard deviation of SD_U was equal to 0.009 (0.001-0.02). Consequently, for only 12 sites (7%) significant difference of IPR+ index values were detected through years, i.e. for a given site, at least for one year, the IPR+ score distribution significantly differed from the 10-year median score (Fig. 2). Ten out of these 12 sites were highly disturbed sites (Gp4-Gp5) and two were slightly disturbed (Gp2). Except for one site, always one year was significantly different from the median score (e.g. Fig. 2c). For only two sites, one year-wise comparison showed a significant difference, i.e. IPR+ index distribution significantly different (e.g. Fig. 2c).

Metrics values were mostly consistent among the years with average correlations ranging from r = 0.65 for N%-Trout (the relative number of individuals of trout juveniles 0+) to r = 0.82 for N%-O2INTOL (the relative number of oxyphilous individuals) (Table 2). Except for N%-Trout (mean SD_U = 0.19, SD_T = 0.14) metric SD_U (0.21-0.35) was twice to three times larger than metric SD_T (0.07-14). The percentage of sites with significant difference through years in metric values was often larger than for IPR+ (Table 2), with the highest number for N%-Trout (43%).

3.2. Influence of human-induced disturbances and longitudinal gradient on IPR+ index and metrics inter-annual variability

IPR+ index inter-annual variability took a wide range of values and was not homogenous among sites (Table 3). SD_T of the index score increased significantly with level of human disturbances (F test, $p<0.001$) (Fig. 3a). Moreover, index inter-annual variability was not significantly different for upstream and downstream sites (F test, $p>0.05$) (Fig. 3b).

Except for S-STTHER and N%-Trout, all the metrics showed increasing temporal variability from undisturbed to highly disturbed conditions (F test, $p<0.05$; Table 3).

By contrast, N%-Trout values were more variable for slightly (mean(SD_T)=0.16) than for highly impacted sites (0.08; F test, $p<0.001$).

Moreover, N%-HINTOL, N%-O2INTOL, N%-RHPAR, S-INTOL, S-LIPAR and S%-O2INTOL were more stable through time in upstream than in downstream sites whereas it was the contrary for N%-Trout (F test, $p<0.01$). Except for S%-LIMNO, S-LIPAR, S-OMNI, S-STTHER and S-TOL, interaction effects between upstream-downstream and P_G gradients were not significant (F tests, $p>0.05$).

4. Discussion

This study was conduct in a validation perspective of a bio-indicator developed to assess French river ecological condition and focus on the acknowledgment of its temporal variability unrelated with human degradation. The final aim was to provide recommendations to users. The inter-annual variability of the IPR+ index and underlying metrics was quantified and analysed while accounting for their predictive uncertainty.

IPR+ scores of the 183 sites appeared to be consistent through time. Uncertainty of the metric values and IPR+ score was generally much larger than the inter-annual variability explaining the few sites showing significant inter-annual variations of IPR+ index and metrics. Finally, IPR+ index was significantly more variable at degraded sites than at less impaired sites but showed comparable inter-annual variations along the longitudinal gradient.

4.1. Inter-annual variability versus uncertainty

In stable human disturbance conditions, over a 10-year period, only few sites showed significant inter-annual difference of IPR+ scores. These results showed that IPR+ index inter-annual variability was generally negligible compare to predictive uncertainty of the reference conditions. This result revealed that in absence of changes in human disturbances, whether IPR+ index scores generally do not vary substantially between years or IPR+ index scores do vary but individual scores are too uncertain to allow score distinction between two years. In addition, the Pearson correlations of the 183 site scores between the 10 years confirmed that the IPR+ index give consistent evaluation over time. The evaluation of the index stability should be conducted over an appropriate temporal scale, which has been argued to be the mean generation time of dominant species of assemblages (Schaeffer *et al.*, 2012). In this study, the data time-frame was limited to 10-year which matched the several years mean life-time of most abundant species of our dataset, e.g. Eurasian minnow (*Phoxinus phoxinus*; several years) and the chub (*Squalius cephalius*; about 10 years) (Keith *et al.*, 2011).

As for IPR+ index, temporal variability and uncertainty of the metrics were variable among sites and were different depending on the metric considered. Nevertheless, metrics SD_T and SD_U were often higher than for IPR+ and Pearson correlations showed that metrics are globally less stable than the index. These two results indicate that for multi-metric index the choice and combination of particular metrics may reduce temporal variability and uncertainty of the index (in the case of IPR+, the 6 most degraded metrics are used to calculate the index; see Marzin *et al., submitted*). This is consistent with the use of several metrics synthesised into an individual index rather than the use of metric alone. It seems that multimetric indices may lessen the temporal variability of metrics related to the environmental stochasticity (Bêche et al. 2006) and accordingly enable a more reliable assessment of stream ecological conditions trough years.

Apart from the relative abundance of trout juvenile (N%-Trout), similar differences between SD_U and SD_T were observed for the metrics used to compute the IPR+ index. With S%-LIMNO and S%-O2INTOL, the N%-Trout, although presenting in average the smallest SD_U, was in average the more variable metrics through time. For, N%-Trout, this result was consistent with the others measures of variability, with weakest Pearsons correlations between years and the largest number of sites with significant differences in metric values. This metric was chosen to fulfil WFD requirements and these differences in temporal variations and uncertainty might be connected to the nature of the metric. Abundance of trout young-of-the-year is known to be really influenced by density-dependence phenomenon (e.g. Jenkins *et al.*, 1999; Kaspersson and Höjesjö, 2009) which could not be taken into account into the model. Consequently, this part of N%-Trout variability remained and stood certainly for a substantial part of the temporal variability. More generally, functional assemblage structure is more stable between years than taxonomic composition and species abundance (Bêche *et al.*, 2006; Fransen *et al.*, 2011). Functional trait metrics are relatively impervious to temporal turnover due to shift of dominance by species with similar traits, but they detect changes in functional structure of community. This result supports the idea that metrics based on biological and ecological functional traits should be preferred to metrics based on taxonomic composition (Marzin *et al.*, 2012; Usseglio-Polatera *et al.*, 2000) as they may contribute to stabilize multimetric indexes.

Even if in absence of changes in human pressures inter-annual variability of the IPR+ index exists and may lead to status class differences, the uncertainty associated with index values appears to overcome the temporal variability leading to rare significant temporal differences. As a result, in the case of high uncertainty, the index will permit to detect strong human impact differences between sites but will not allow the interpretation of slight changes. This result emphasizes the importance of considering class confidences and uncertainty to interpret biological indicator scores and be able to differentiate noise from responses to human disturbance.

4.2 Influences of human disturbances and longitudinal gradient on temporal variability

Despite the wide SD_T range of the 183 sites, our findings were in agreement with several studies analysing bio-indicators temporal variability and showing that in degraded situations indicators were less stable over time than in slightly disturbed conditions (Fore *et al.*, 1994; Ross *et al.*, 1985). This is inconsistent with Pyron *et al.* (1998) results on the temporal variability of IBI. The type of streams considered could be the reason as these authors focused on very large streams whereas our data set mostly included medium size streams (median = 224 km^2).

Apart from N%-Trout that showed the opposed pattern and S-STTHER that did not show significant response to P_G, all the metrics presented similar patterns. We assumed that metric inter-annual variability partly resulted from functional structure changes due to human disturbances. The inter-annual variation of community, in terms of functional structure, tends to increase where the underlying disturbances were more severe (Collier, 2008). Also, recent studies on fish community suggest that species loss limits the stabilizing forces of compensatory dynamics (e.g. shift of dominance of species with similar traits) that mitigates variation in aggregate community properties (Franssen *et al.*, 2011; Schaeffer *et al.*, 2012). Subsequently, human disturbances leading to losses of species may destabilize communities and the metrics used to describe their characteristics. The lack of inter-annual stability of the index might give an additional evidence of the strong degradation of the system.

The reverse pattern observed for the N%-Trout, might be explained by the more variable nature of compositional metrics and by the fact that trout juvenile could not stand strong degradation. Consequently, it is quite probable that the larger range of responses of this metric will occur at the beginning of the human disturbance gradient (Gp1-Gp3) and decrease sharply to more stable values close to zero in highly disturbed conditions (Gp4-Gp5.

In addition, some of the metrics were more stable through time in upstream than in downstream sites, whereas IPR+ stability was not significantly different. This result is opposed to previous works (Horwitz 1978; Schlosser 1982; Oberdorff et al. 2001) suggesting that fish community are more stable downstream than upstream. The instability of the assemblage might reflect the difficulty to achieve representative samples in larger rivers. Since the abundance and number of species are generally larger downstream (Huet 1954; Horwitz 1978; Oberdorff & Porcher 1992), the sampling effort might not always be sufficient to represent the whole community present in the river, resulting in increasing composition variability downstream. Furthermore, this could be accentuated by the sampling techniques used to sampled downstream sites. In most of the downstream sites, assemblage composition was assessed by boat-sampling which has been shown to be more variable between years than backpack electrofishing used in wadable stream (Meador and McIntyre, 2003). Those results are consistent with Collier's work on macroinvertebrate communities (2008) showing relationships between metric stability and landscape factors but no relationship between landscape factors and the Macroinvertebrate Community Index stability. As for scores, the aggregation of metric into a synthetic index seems to enable score temporal variability to be unrelated to environmental gradients, which is if major importance for water managers when evaluating stream conditions.

For the IPR+ index and most of the metrics, interaction effects between upstream-downstream and P_G gradients were not significant indicating that the effects of the pressure intensity on temporal variability were comparable in upstream and downstream sites. To balance the data set, human disturbance intensity and longitudinal gradients were cut into coarse categories that certainly blur a more complex reality. Nevertheless, it is rare to acquire datasets robust enough to test simultaneously the influence of multiple factors and the aim here was to detect the main factors and patterns influencing the inter-annual variability of IPR+ index and metrics.

Our results suggest that beside its direct impact on fish assemblage structure, an increase in human disturbance intensity weakens the temporal stability of fish assemblage functional structure and consequently affects the temporal stability of bio-assessment indicators relying on functional metrics.

4.3 Multi-metric index, temporal variability and implications for bio-assessment

Since they undoubtedly affect the bio-assessment reliability, it is of prime importance to identify the main sources of index temporal variability and uncertainty for future improvements. The sources are multiple but several have been recognized as primordial in this study. First, in the case of the IPR+ index, models used to predict the expected values of descriptors of the assemblage structure relied on environmental variables including climatic variables over the 10-year previous to sampling and thus accounted for mid-term reference condition changes linked to global climatic evolutions. Nevertheless, exceptional annual climatic or hydrological events known to affect fish assemblage structure (Schaeffer *et al.*, 2012), such as flooding or drought, were not integrated in the models and might have reduced inter-annual variation and predictive uncertainty observed for IPR+ index and metrics (Mazor *et al.*, 2009; Linke *et al.*, 1999). This reflection also strengthens the need to sample reference sites through time to account for reference conditions shift and global changes for future bio-assessments (Logez and Pont, 2012).

Second, temporal changes in assemblage composition and structure could not be only a direct and predictable consequence of environmental changes but also be a result of biological or stochastic processes such as recruitment (as illustrated by N%-Trout) or extinctions. The natural temporal dynamism of species assemblages may be a source of temporal variability and predictive uncertainty and consequently could bias bio-assessments. As noted previously, metrics based on ecological and biological traits seem generally less sensitive to natural biological processes and should be preferred to metrics based on species composition.

One of the most important sources of temporal variability and uncertainty of index and metrics is certainly linked to the sampling error and different sampling efforts between years. This last type could be reduced by using standardised protocol and improving crew experience and acquaintance with sampling methods (Hardin and Connor, 1992; Benejam *et al.*, 2012).

Finally, the use of predictive multimetric index that considers environmental conditions to predict metric values in minimally disturb conditions allows index values and variability to free from landscape factors. Moreover, the temporal variability that could be observed in absence of degradations is overcome by the uncertainty associated with reference condition establishment which highlights the necessity for assessing uncertainty of bio-assessment tools.

5. Conclusion

As bio-indicators are devoted to evaluate the ecological quality of water bodies, their uncertainty and uncontrolled variability have crucial implications for water management decisions and need to be acknowledged. In absence of human disturbance changes, evaluations of the 183 French sites given by IPR+ scores were globally consistent through time. Predictive uncertainty related to the establishment of IPR+ index and metrics' values in reference conditions was generally much larger than the inter-annual variability. This can explain the weak numbers of sites showing significant inter-annual variations. In addition, IPR+ index showed comparable inter-annual variations along the upstream-downstream gradient and was significantly less stable at degraded sites than at more pristine sites.

Therefore, in addition to weak index scores, high inter-annual variability of the index might be an evidence of strong impacts of human disturbances on river fish communities.
Finally, for fish assemblages, functional metrics were less sensitive to natural temporal dynamics of assemblages than compositional metrics and thus give a more accurate image of the impact of human-induced disturbances on fish assemblages. These results confirm the advantages of multi-metrics index based on ecological and biological functional metrics for the assessment of river ecological health and endorse the need for sampling reference sites through time to account for reference conditions shift and global changes for future bio-assessments.

Aknowledgements

This work has been supported by the collaborative project between Irstea and the French Electric Company (EDF): "HYNES". We would like to thank Météo-France and the Hydrology Research Team, Irstea Antony for their support on climatic data.

References

Bady, P., Doledec, S., Dumont, B. & Fruget, J.F. (2004) Multiple co-inertia analysis: a tool for assessing synchrony in the temporal variability of aquatic communities. Comptes Rendus Biologies, 327(1), 29-36.

Bady, P., Doledec, S., Fesl, C., Gayraud, S., Bacchi, M. & Scholl, F. (2005) Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. Freshwater Biology, 50(1), 159-73.

Bailey, R.C., Kennedy, M.G., Dervish, M.Z. & Taylor R.M. (1998) Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrates communities in Yukon streams. Freshwater Biology, 39, 765-774.

Bêche, L.A., McElravy, E.P. & Resh, V.H. (2006) Long-term seasonal variation in the biological traits of benthic-macroinvertebrates in two Mediterranean-climate streams in California, U.S.A. Freshwater Biology, 51(1), 56-75.

Benejam, L., Alcaraz, C., Benito, J., Caiola, N., Casals, F., Maceda-Veiga, A., de Sostoa, A. & Garcia-Berthou, E. (2012) Fish catchability and comparison of four electrofishing crews in Mediterranean streams. Fisheries Research, 123, 9-15.

Bunn, S.E. & Davies, P.M. (2000) Biological processes in running waters and their implications for the assessment of ecological integrity. Hydrobiologia, 422-423, 61-70.

Carlisle, D.M. & Clements, W.H. (1999) Sensitivity and variability of metrics used in biological assessments of running waters. Environmental Toxicology and Chemistry, 18(2), 285-91.

Clarke, R.T. & Hering, D. (2006) Errors and uncertainty in bioassessment methods - Major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia, 566(1), 433-39.

Collier, K. J. (2008) Temporal patterns in the stability, persistence and condition of stream macroinvertebrate communities: relationships with catchment land-use and regional climate. Freshwater Biology, 53(3), 603-616.

Fore, L.S., Karr, J.R. & Conquest, L.L. (1994) Statistical Properties of an Index of Biological Integrity Used to Evaluate Water-Resources. Canadian Journal of Fisheries and Aquatic Sciences, 51(5), 1077-87.

Franquet, E., Dolédec, S. & Chessel, D. (1995) Using multiviriate analyses for separating spatial and temporal effects within species-environment relationships. Hydrobiologia, 300-301(1), 425-31.

Franssen, N.R., Tobler, M. & Gido, K.B. (2011) Annual variation of community biomass is lower in more diverse stream fish communities. Oikos, 120(4), 582-90.

Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4(4), 379-91.

Growns, I., Astles, K. & Gehrke, P. (2006) Multiscale spatial and small-scale temporal variation in the composition of riverine fish communities. Environmental Monitoring and Assessment, 114(1-3), 553-71.

Hardin, S., and Connor L. L. (1992) Variability of electrofishing crew efficiency, and sampling requirements for estimating reliable catch rates. North American Journal of Fisheries Management 12, 612–617.

Hartigan, J.A. & Wong, M.A. (1979) A K-means clustering algorithm. Applied Statistics, 28,100-108.

Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K. & Verdonschot, P.F.M. (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: A comparative metric-based analysis of organism response to stress. *Freshwater Biology*, **51**, 1757-1785.

Horwitz, R.J. (1978) Temporal Variability Patterns and the Distributional Patterns of Stream Fishes. Ecological Monographs, 48(3), 307-21.

Huang, G.H. & Chang, N.B. (2003) Perspectives of Environmental Informatics and Systems Analysis. Journal of Environmental Informatics, 1(1), 1-6.

Huet, M. (1954) Biologie, profils en long et en travers des eaux courantes. Bulletin Français de Pisciculture, 175, 41-53.

Hughes, R.M., Howlin, S. & Kaufmann, P.R. (2004) A biointegrity index (IBI) for coldwater streams of Western Oregon and Washington. Transactions of the American Fisheries Society, 133, 1497-1515.

Jenkins, T.M., Diehl, S., Kratz, K.W. & Cooper, S.D. (1999) Effects of population density on individual growth of brown trout in streams. Ecology, 80(3), 941-56.

Kaspersson, R. & Hojesjo, J. (2009) Density-dependent growth rate in an age-structured population: a field study on stream-dwelling brown trout Salmo trutta. Journal of Fish Biology, 74(10), 2196-215.

Keith, P., Persat, H., Feuteun, E. & Allardi, J. (2011) Les poissons d'eau douce de France. Biotope, Mèze; Museum national d'histoire naturelle, Paris (collection Inventaires et biodiversité).

Linke, S., Bailey, R.C. & Schwindt, J. (1999) Temporal variability of stream bioassessments using benthic macroinvertebrates. Freshwater Biology, 42(3), 575-84.

Logez, M. & D. Pont. (2012) Global warming and potential shift in reference conditions: the case of functional fish-based metrics. Hydrobiologia.

Logez, M., Bady, P., Melcher, A. & Pont, D. (2012) A continental-scale analysis of fish assemblage unctional structure in European rivers. Ecography. doi: 10.1111/j.1600-0587.2012.07447.x

Marzin, A., Archaimbault, V., Belliard, J., Chauvin, C., Delmas, F. & Pont, D. (2012) Ecological assessment of running waters: Do macrophytes, macroinvertebrates, diatoms and fish show similar responses to human pressures? *Ecological Indicators*, **23**, 56-65.

Marzin, A., Belliard J., Delaigue, O., Logez M., Pont D. (Submitted to Canadian Journal of fisheries sciences). Uncertainty associated with river health assessment in a varying environment: the case of a predictive fish-based index in France.

Mazor, R.D., Purcell, A.H. & Resh, V.H. (2009) Long-Term Variability in Bioassessments: A Twenty-Year Study from Two Northern California Streams. Environmental Management, 43(6), 1269-86.

Meador, M.R. & McIntyre, J.P. (2003) Effects of Electrofishing Gear Type on Spatial and Temporal Variability in Fish Community Sampling. Transactions of the American Fisheries Society, 132(4), 709-16.

Oberdorff, T. & Porcher, J.P. (1992) Fish assemblage structure in Brittany streams (France). Aquatic Living Resources, 5, 215-223.

Oberdorff, T., Pont, D., Hugueny, B. & Chessel, D. (2001) A probabilistic model characterizing fish assemblages of French rivers: A framework for environmental assessment. Freshwater Biology, 46(3), 399-415.

Oberdorff, T., Pont, D., Hugueny, B. & Porcher, J.P. (2002) Development and validation of a fish-based index for the assessment of 'river health' in France. Freshwater Biology, 47(9), 1720-34.

Ostermiller, J.D. & Hawkins, C.P. (2004) Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society, 23(2), 363-82.

Palmer, M.A., Bernhardt, E.S., Allan, J.D., Lake, P.S., Alexander, G., Brooks, S., Carr, J., Clayton, S., Dahm, C.N., Shah, J.F., Galat, D.L., Loss, S.G., Goodwin, P., Hart, D.D., Hassett, B., Jenkinson, R., Kondolf, G.M., Lave, R., Meyer, J.L., O'Donnell, T.K., Pagano, L. & Sudduth, E. (2005) Standards for ecologically successful river restoration. Journal of Applied Ecology, 42(2), 208-17.

Pont, D., Hugueny, B. & Oberdorff, T. (2005) Modelling habitat requirement of European fishes : do species have similar responses to local and regional environmental constraints? Canadian journal of fisheries and aquatic sciences, 62, 163-173.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N. & Schmutz, S. (2006) Assessing river biotic condition at the continental scale: a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology*, **43**, 70-80.

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

Pyron, M., Lauer, T., LeBlanc, D., Weitzel, D. & Gammon, J. (2008) Temporal and spatial variation in an index of biological integrity for the middle Wabash River, Indiana. Hydrobiologia, 600(1), 205-14.

Ricklefs, R.E. & Schluter, D. (1993). Species diversity in ecological communities: historical and geographical perspectives. In Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press.
Ross, S.T., Matthews, W.J. & Echelle, A.A. (1985) Persistence of Stream Fish Assemblages - Effects of Environmental-Change. American Naturalist, 126(1), 24-40.
Schaefer, J.F., Clark, S.R. & Warren, M.L. (2012) Diversity and stability in Mississippi stream fish assemblages. Freshwater Science, 31(3), 882-94.
Schlosser, I.J. (1982) Fish Community Structure and Function Along 2 Habitat Gradients in a Headwater Stream. Ecological Monographs, 52(4), 395-414.
Spiegelhalter, D., Thomas, A., Best, N. & Lunn, D. (2003) WinBUGS User Manual. Version 1.4, http://www.mrc-bsu.cam.ac.uk/bugs.
Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K. & Norris, R.H. (2006) Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications, 16(4), 1267-76.
Tenenhaus, M. & Young, F.W. (1985) An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis ans other methods for quantifying categorical multivariate data. Psychometrika, 50 (1), 91-119.
Usseglio-Polatera, P., Bournaud, M., Richoux, P. & Tachet, H., 2000. Biomonitoring through biological traits of benthic macroinvertebrates: How to use species trait databases? Hydrobiologia 422-423, 153-162.
Vidal, J.P., Martin, E., Franchistéguy, L., Baillon, M. & Soubeyroux, J.M. (2009) A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30, 1627-1644.
Wang, L.Z., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. & Stewart, P.M. (2003) Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences, 60(5), 491-505.
Willby N., Birk S. & Bonne W. (2010) IC Guidance Annex V: Definition of comparability criteria for setting class boundaries, 22 p.
Working Group Ecostat (2009) Implementation strategy for the Water Framework Directive (2000/60/Ec) Guidance document no. 14. Guidance document on the intercalibration process 2008-2011, 55 p.

Table 1 Description of the human pressures degrading the 183 French sites.

Reach-scale human disturbances	Level of disturbances (number of sites)
Barrier to sea	No (18) / Partial (20) / Yes (145)
Upstream barrier	No (101) / Partial (42) / Yes (40)
Downstream barrier	No (92) / Partial (44) / Yes (47)
Hydrological regime modified	No (79) / Slight (57) / Int. (29) / High (18)
Hydropeaking	No (156) / Slight (11) / Moderate to High (16)
By-pass channel/diversion	No (170) / Yes (13)
Water abstraction	No (79) / Slight (72) / Int. (13) / High (19)
Presence of an impoundment	No (145) /Partial (21) / Yes (17)
Velocity increased	No (175) / Yes (8)
Artificial embankment	No (137) / Slight (29) / Moderate to High (17)
Riparian vegetation modified	No (99) / Slight (36) / Int. (40) / High (8)
Sedimentation	No (71) / Slight (54) / Int. (44) / High (14)
Channel form modified	No (134) /Int. (28) / High (21)
Cross-section modified	No (121) /Int. (33) / High (29)
Instream habitat modified	No (121) /Int. (52) / High (10)
MorphoTransSolid = incision	No (120) /Int. (49) / High (14)
Diking	No (155) /Int. (15) / High (13)
Water temperature modified	No (137) /Warming (39) / Cooling (7)
Toxic pollution	No (99) / Slight (49) / Int. (23) / High (12)
Organic pollution	No (94) / Slight (63) / Int. (19) / High (7)
Nutrient pollution	No (79) / Slight (57) / Int. (38) / High (9)
Eutrophication	No (120) / Slight (34) / Moderate to High (29)
Upstream lake	No (117) /Partial (36) / Yes (30)

Table 2 Temporal variability (inter-annual standard deviation, SD_T) and uncertainty (PDFs standard deviation, SD_U) of the IPR+ index and metrics (Mean (Min-Max)). Average inter-annual Pearson correlations of the IPR+ scores and metric values. %-Temp-var-Median = the percentage of sites where IPR+ scores/Metric values (1998-2007) were at least for one year, significantly different from the median score (no overlap of the PDF 2.5% and 97.5% percentiles with the median value). %-Temp-var-Pairs = the percentage of sites with at least one of the pair-year comparison showing significant distribution differences.

	SD_T	SD_U	Pearson r	%-Temp-var-Median	%-Temp-var-Pairs
IPR+	***0.07 (0.01-0.17)***	***0.14 (0.04-0.17)***	***0.82 (0.74-0.88)***	***7%***	***1%***
N%-HINTOL	*0.08 (0.002-0.29)*	*0.22 (0.01-0.3)*	*0.7 (0.56-0.81)*	*7%*	*3%*
N%-O2INTOL	*0.07 (0.002-0.32)*	*0.21 (0.04-0.26)*	*0.82 (0.75-0.9)*	*9%*	*3%*
N%-RHPAR	*0.06 (0.002-0.27)*	*0.24 (0-0.36)*	*0.68 (0.5-0.81)*	*8%*	*4%*
N%-Trout	*0.14 (0.0005-0.27)*	*0.19 (0.01-0.27)*	*0.65 (0.44-0.88)*	*43%*	*11%*
S%-INTOL	*0.07 (0.003-0.29)*	*0.3 (0.14-0.43)*	*0.69 (0.54-0.81)*	*0%*	*0%*
S%-LIMNO	*0.09 (3e-8-0.4)*	*0.27 (0-0.42)*	*0.78 (0.7-0.85)*	*14%*	*13%*
S%-O2INTOL	*0.1 (0.0005-0.36)*	*0.26 (0.04-0.42)*	*0.75 (0.66-0.83)*	*14%*	*10%*
S-LIPAR	*0.07 (0.004-0.23)*	*0.32 (0.07-0.45)*	*0.78 (0.71-0.86)*	*1%*	*1%*
S-OMNI	*0.08 (0.003-0.24)*	*0.3 (0.07-0.43)*	*0.79 (0.68-0.88)*	*2%*	*0%*
S-STTHER	*0.07 (0.003-0.28)*	*0.35 (0.17-0.44)*	*0.68 (0.5-0.86)*	*1%*	*0%*
S-TOL	*0.09 (0.002-0.27)*	*0.3 (0.03-0.43)*	*0.74 (0.65-0.84)*	*3%*	*1%*

Table 3 Mean inter-annual variability of the IPR+ index and metrics. Calculated for the sites slightly (Gp1-Gp2), intermediately (Gp3), and highly disturbed by human activities (Gp4-Gp5), and upstream and downstream sites (i.e. Trout-Grayling and Barbel-Bream zones; Huet, 1954). Effects of the longitudinal gradient (Up-Down) and of human-induced disturbances (P_G) were tested performing ANOVA procedures (*p*-value: *** ≤0.001, ** ≤0.01 and * ≤0.05 and *NS* for non-significant tests).

	P_G gradient			Longitudinal gradient		ANOVA test		
	Gp1-Gp2	Gp3	Gp4-Gp5	Upstream	Downstream	P_G	Up-Down	Interations
IPR+	*0.06*	*0.07*	*0.08*	*0.06*	*0.08*	***	*ns*	*ns*
N%-HINTOL	*0.05*	*0.08*	*0.12*	*0.06*	*0.11*	***	**	*ns*
N%-O2INTOL	*0.03*	*0.08*	*0.1*	*0.04*	*0.1*	***	***	*ns*
N%-RHPAR	*0.03*	*0.05*	*0.09*	*0.04*	*0.09*	***	***	*ns*
N%-Trout	*0.16*	*0.12*	*0.08*	*0.14*	-	***	-	-
S%-INTOL	*0.06*	*0.07*	*0.1*	*0.06*	*0.09*	***	**	*ns*
S%-LIMNO	*0.07*	*0.1*	*0.1*	*0.08*	*0.1*	**	*ns*	*
S%-O2INTOL	*0.06*	*0.1*	*0.13*	*0.06*	*0.13*	***	***	*ns*
S-LIPAR	*0.05*	*0.08*	*0.09*	*0.06*	*0.09*	***	***	*
S-OMNI	*0.06*	*0.09*	*0.09*	*0.07*	*0.09*	*	*ns*	*
S-STTHER	*0.06*	*0.07*	*0.08*	*0.06*	*0.08*	*ns*	*ns*	***
S-TOL	*0.07*	*0.1*	*0.1*	*0.08*	*0.1*	**	*ns*	**

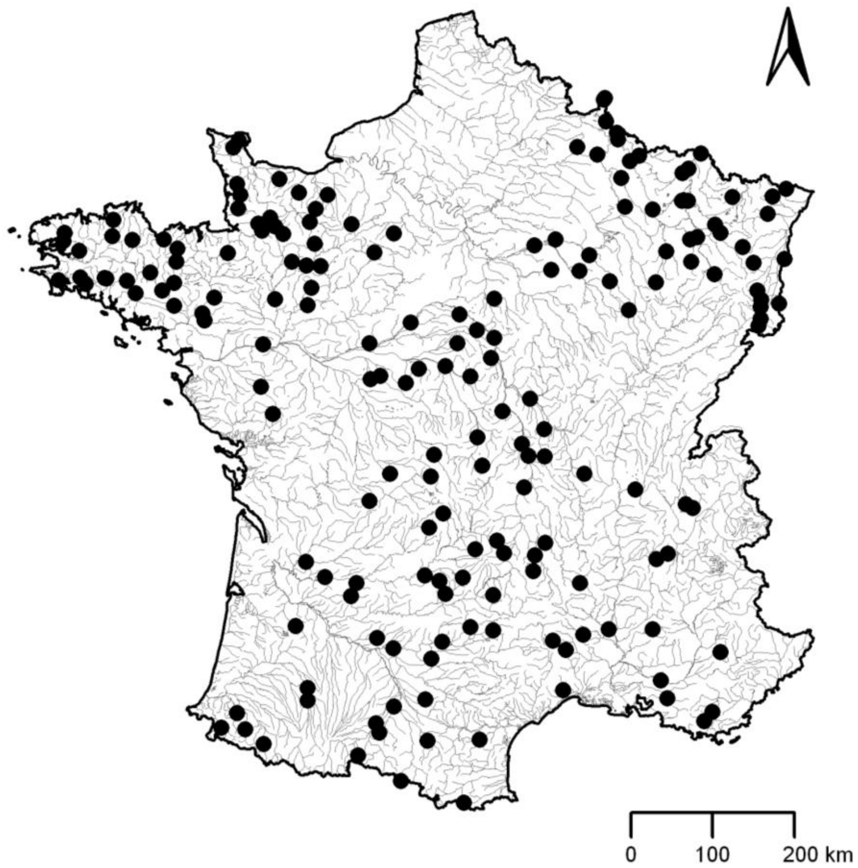


Fig. 1 Locations of the 183 French sites

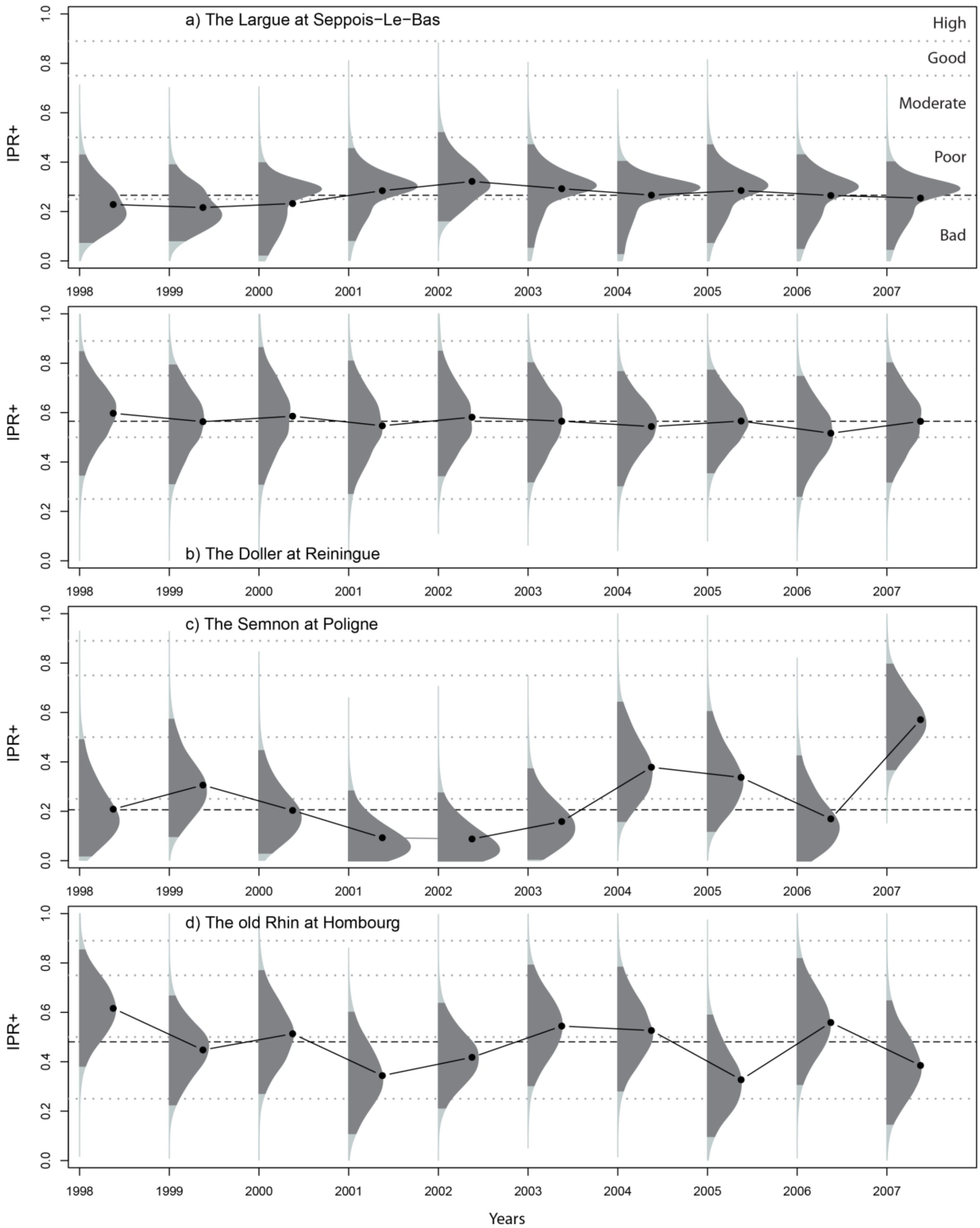


Fig. 2 Examples of IPR+ index 10-year inter-annual variability (SDT) and uncertainty (SDU). Black dashed horizontal line, 10-year median of the IPR+ scores. Black dots, IPR+ score, i.e. mean of the PDFs. Dark grey area, 95% probability of the IPR+ score. Light grey area, the other 5%. Grey dotted lines, thresholds of the five ecological status classes (high, good, moderate, poor, bad) defined in agreement with European intercalibration rules (Working Group Ecostat 2009; Willby *et al.*, 2010)

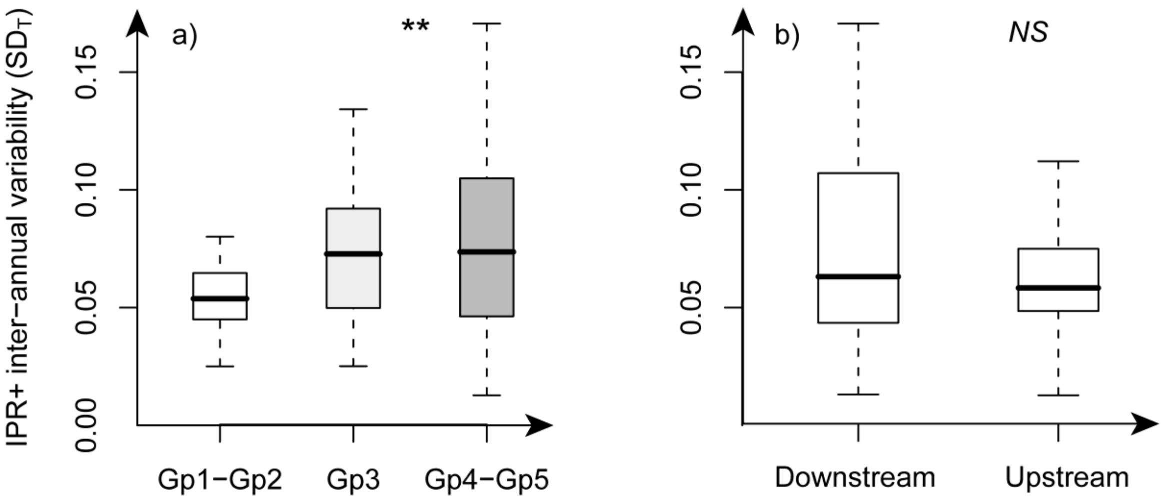


Fig. 3 IPR+ inter-annual variability (a, b) relationships with human-induced disturbance intensity (P_G) and longitudinal gradient (*p*-value: ** <0.01 and *NS* for non-significant tests)

P5

Restoration by Removal of Weirs and Dams (< 5m Height)

Chapter in "From Natural to Degraded Rivers and Back Again: A Test of Restoration Ecology Theory and Practice."

Feld, C.K., S. Birk, D.C.Bradley, D. Hering, J. Kail, **A. Marzin**, A. Melcher, D. Nemitz, M.L.Pedersen, F. Pletterbauer, D. Pont, P.F.M. Verdonschot & N. Friberg (2011)

Advance in Ecological Research Vol. 44, 120-209.

From Natural to Degraded Rivers and Back Again: A Test of Restoration Ecology Theory and Practice

CHRISTIAN K. FELD, SEBASTIAN BIRK, DAVID C. BRADLEY, DANIEL HERING, JOCHEM KAIL, ANAHITA MARZIN, ANDREAS MELCHER, DIRK NEMITZ, MORTEN L. PEDERSEN, FLORIAN PLETTERBAUER, DIDIER PONT, PIET F.M. VERDONSCHOT AND NIKOLAI FRIBERG

ADVANCES IN ECOLOGICAL RESEARCH VOL. 44 0065-2504/11 $35.00

DOI: 10.1016/B978-0-12-374794-5.00003-1

SUMMARY

Extensive degradation of ecosystems, combined with the increasing demands placed on the goods and services they provide, is a major driver of biodiversity loss on a global scale. In particular, the severe degradation of large rivers, their catchments, floodplains and lower estuarine reaches has been ongoing for many centuries, and the consequences are evident across Europe. River restoration is a relatively recent tool that has been brought to bear in attempts to reverse the effects of habitat simplification and ecosystem degradation, with a surge of projects undertaken in the 1990s in Europe and elsewhere, mainly North America. Here, we focus on restoration of the physical properties (e.g. substrate composition, bank and bed structure) of river ecosystems to ascertain what has, and what has not, been learned over the last 20 years.

First, we focus on three common types of restoration measures—riparian buffer management, instream mesohabitat enhancement and the removal of weirs and small dams—to provide a structured overview of the literature. We distinguish between abiotic effects of restoration (e.g. increasing habitat diversity) and biological recovery (e.g. responses of algae, macrophytes, macroinvertebrates and fishes).

We then addressed four major questions: (i) Which organisms show clear recovery after restoration? (ii) Is there evidence for qualitative linkages between restoration and recovery? (iii) What is the timescale of recovery? and (iv) What are the reasons, if restoration fails?

Overall, riparian buffer zones reduced fine sediment entry, and nutrient and pesticide inflows, and positive effects on stream organisms were evident. Buffer width and length were key: 5–30 m width and >1 km length were most effective. The introduction of large woody debris, boulders and gravel were the most commonly used restoration measures, but the potential positive effects of such local habitat enhancement schemes were often likely to be swamped by larger-scale geomorphological and physico-chemical effects. Studies demonstrating long-term biological recovery due to habitat enhancement were notable by their absence. In contrast, weir removal can have clear beneficial effects, although biological recovery might lag behind for several years, as huge amounts of fine sediment may have accumulated upstream of the former barrier.

Three Danish restoration schemes are provided as focal case studies to supplement the literature review and largely supported our findings. While the large-scale re-meandering and re-establishment of water levels at River Skjern resulted in significant recovery of riverine biota, habitat enhancement schemes at smaller-scales in other rivers were largely ineffective and failed to show long-term recovery.

The general lack of knowledge derived from integrated, well-designed and long-term restoration schemes is striking, and we present a conceptual framework to help address this problem. The framework was applied to the three restoration types included in our study and highlights recurrent cause–effect chains, that is, commonly observed relationships of restoration measures (cause) and their effects on abiotic and biotic conditions (effect). Such conceptual models can provide useful new tools for devising more effective river restoration, and for identifying avenues for future research in restoration ecology in general.

V. RESTORATION BY REMOVAL OF WEIRS AND DAMS (<5 M HEIGHT)

Among 31 restoration studies (and five additional general ecological papers), the majority examined the effects of weir removal at North American streams, and only one restoration study (Tszydel *et al.*, 2009) and two reviews originated from Europe (de Leaniz, 2008; Schmitt, 2005). Most of the studies sampled several stretches per measure, each of which was several hundred metres long and together spanned stream sections of one to a few kilometres in length. The comparison of conditions before and after restoration was common to all restoration studies, while roughly half of the studies applied a full BACI design. The ecological effects of weir removal were comprehensively reviewed by Bednarek (2001), who also presented a series of case studies to underpin their review with empirical data.

Bunn and Arthington (2002) stressed the role of flow as a major determinant of physical habitat in streams which, in turn, was a major determinant of biotic composition. Acreman and Dunbar (2004) referred to the flow regime required in a river to achieve desired ecological objectives, that is, the 'environmental flow', and multiple elements of which are important (Poff, 1997). Low flows provide minimum habitat for resident species, medium flows sort river sediments and stimulate fish migration and spawning, and floods maintain channel structure and allow movement between floodplain habitats (Acreman and Dunbar, 2004). Occasional floods reconnect the aquatic and riparian habitat (Shuman, 1995), and backwaters are refilled. Fine materials (e.g. sand, silt, mud) erode and uncover coarser substrata (e.g. gravel, pebble and cobbles), which enhances the overall habitat diversity (Born *et al.*, 1996; Kanehl *et al.*, 1997) and dissolved oxygen and water quality improve (Hill *et al.*, 1993). Bednarek (2001), however, also referred to some negative effects, such as the downstream transport of contaminated sediments or the overall abrasive effect of fine sediment movement, although these adverse effects are typically short-term, whereas overall improvement is more likely to occur in the longer term.

The changing abiotic conditions may improve biodiversity and reproduction of fish: the spawning grounds for salmonid species increase (Iversen *et al.*, 1993), while fish passage is more likely because of the restored longitudinal connectivity. Migratory fishes depend on the re-establishment of hydrological connectivity, which was a common key argument for restoration (Iversen *et al.*, 1993; Poff, 1997; but see also de Leaniz, 2008 for a more recent summary of findings during the past decade). The maintenance not only of the longitudinal but also of the lateral connectivity with the floodplain is essential to the viability of populations of many riverine species (Bunn and Arthington, 2002; Bushaw-Newton *et al.*, 2002; Maloney *et al.*, 2008).

Stanley and Doyle (2003) suggested weir removals may be best considered as ecological disturbances, as removal of small dams generally results in the transformation from lentic to lotic river systems upstream, leading to the reservoir sediment release and a pulse downstream, which could cause short-term reductions in productivity and possibly diversity (Bednarek, 2001). In addition, effects of restoration could be very variable depending on the hydrologic nature of the river (Chaplin, 2003). As a result, the effectiveness of a dam removal, that is, the recovery of a river from the induced disturbance, is likely to vary widely among systems and to be contingent upon both temporal and spatial scales of the restoration scheme.

The literature provided little information on the effectiveness of such restorations: it was rarely measured and the judging criteria were usually vague. In most cases, negative impacts of weir removal were rather short-term (e.g. increase in suspended sediments) while the assumed beneficial changes were likely to act in the longer term (e.g. increase in flow diversity, connectivity); the natural free-flowing state of the river was always regained whereas recovery of the biota following this habitat shift was more uncertain. Five consistent effects of weir removal were identified relating to: morphology (width and depth); substrate particle size (and gravel bars); flow diversity (and turbidity); temperature and connectivity (Figure 11). Several studies

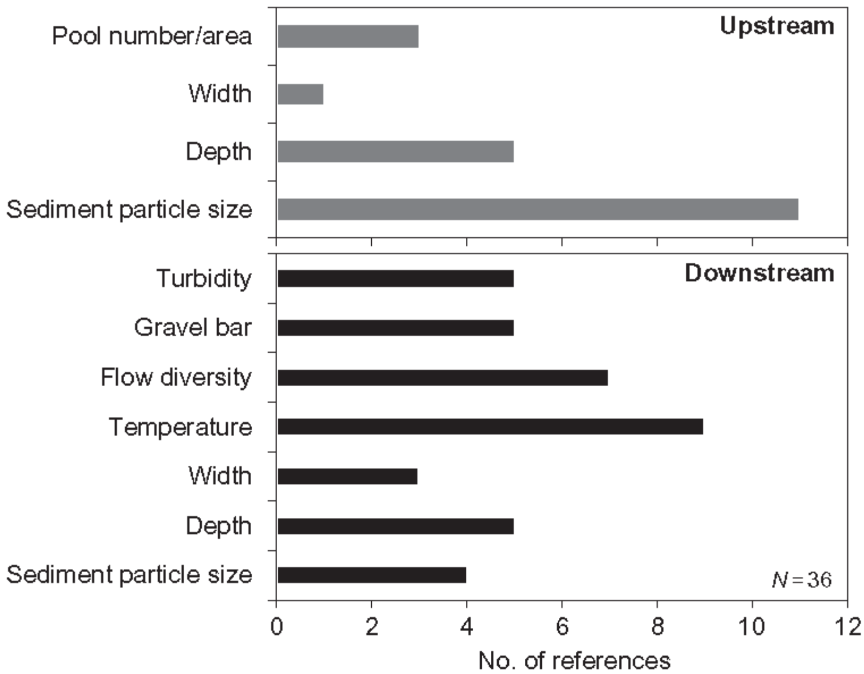


Figure 11 Ranking of the most important environmental state variables based on the linkages (arrow thickness in Figure A3), derived from 36 references on weir removal.

(e.g. Chaplin, 2003; Cheng and Granata, 2007) found an increase in sediment particle size upstream and a decrease downstream of the former weir and other studies commonly reported an increase in flow diversity upstream (e.g. Hill *et al.*, 1993), decreased water temperature upstream (e.g. Hill *et al.*, 1993; Kanehl *et al.*, 1997) and restoration of the hydro-ecological connectivity (e.g. Gregory *et al.*, 2002; Poff, 1997).

In summary, the removal of weirs and small dams re-establishes natural physical river characteristics with some evidence of positive effects for the instream fauna, in particular for migrating fishes for whom it is arguably a general prerequisite for restoring their populations. The full beneficial effects of weir removal (e.g. Gregory *et al.*, 2002), however, may take decades to be manifested, while some adverse effects are likely to dominate in the short-term due to the mobilisation of fine sediments (e.g. Orr *et al.*, 2006; Pollard and Reed, 2004; Thomson *et al.*, 2005) accumulated in the former stagnant section (i.e. upstream of the weir). The deposition of this material further downstream on gravel areas, included artificial riffles introduced in other restoration schemes, may limit the availability of spawning habitat for fish in the short-term (i.e. up to several years), so spatial and temporal aspects of this form of restoration are again important (e.g. weirs should not be removed close to the spawning season).

A. Which Organism Groups and Group Attributes Have Shown Evidence of Recovery After Restoration?

The biological impact of weir removal was studied most often for benthic macroinvertebrates (83% of all references), and, to a lesser extent, aquatic macrophytes and fishes (58% and 50%, respectively); phytobenthos was rarely considered (Figure 12). Most commonly, community composition and abundance measures were used to indicate changes due to restoration, irrespective of the organism group. Some papers also considered effects on community functional metrics such as benthic macroinvertebrate feeding habits (e.g. Maloney *et al.*, 2008) or fish growth (e.g. Harvey and Stewart, 1991; Schlosser, 1982). Twelve papers studied the effects of weir removal on sensitive and tolerant benthic macroinvertebrates (mainly EPT taxa: Ephemeroptera–Plecoptera–Trichoptera) and the effects of water quality improvement, such as the abatement of turbidity and oxygen enrichment (Bushaw-Newton *et al.*, 2002; Orr *et al.*, 2006). Changes in the macrophyte community were most often associated with changes in channel morphometry (depth, width) and connectivity (e.g. Shafroth *et al.*, 2002), while the composition and abundance of benthic algae responded to changes in sediment size and turbidity (e.g. Baattrup-Pedersen and Rijs, 1999; Orr *et al.*, 2006).

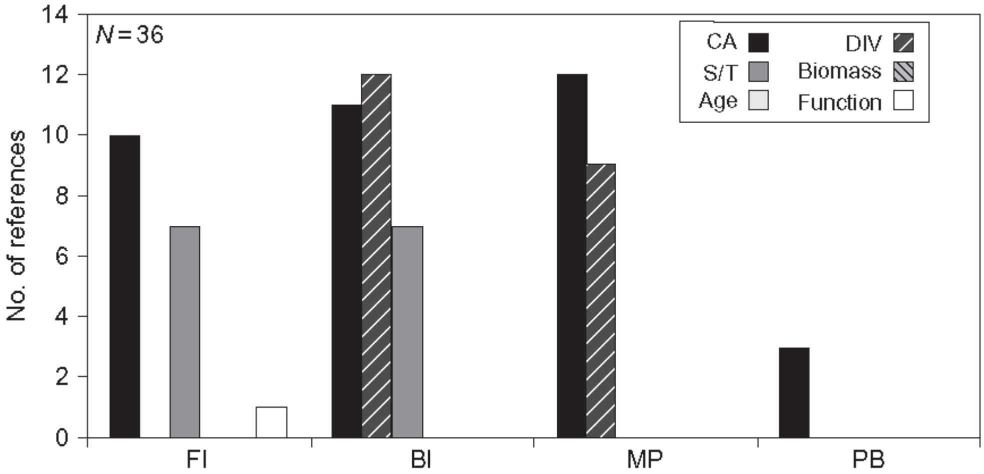


Figure 12 Weir removal: number of references addressing the community attributes composition/abundance (C/A), sensitivity/tolerance (S/T), age structure (Age), diversity (Div), biomass and function of fish (FI), benthic macroinvertebrates (BI), macrophytes (MP) and phytobenthos (PB). As a study may refer to more than one community attribute, the overall number of references exceeds the number of 36 restoration references reviewed.

B. Was There Evidence for Strong Qualitative or Quantitative Linkages?

All restoration studies provided qualitative analyses (Table 3). Nonetheless, sound statistical approaches (ANOVA, ordination) were frequently used to detect and identify patterns of biological impact (e.g. Bushaw-Newton *et al.*, 2002; Pollard and Reed, 2004; Thomson *et al.*, 2005). Cheng and Granata (2007) showed that following removal of a dam, bed deposition and scouring caused a 30% decrease in bed slope and a 40% decrease in bed material size downstream, compared with pre-removal conditions. These impacts reflect the similarly 'strong' linkages reported by Hill *et al.* (1993), Bushaw-Newton *et al.* (2002) and Stanley *et al.* (2002), who revealed a consistent decrease in water temperature upstream, leading to an increase of dissolved oxygen conditions that are likely to favour certain benthic macroinvertebrate and fish species.

In summary, considerable effort has been devoted to investigating the effects of weir removal on riverine systems during the past decade, although there remains a general lack of quantitative results that might help elucidate the mechanistic relationships and provide the means for predictions of the impact of weir removal not only on community recovery but also on community status and recovery further downstream.

Table 3 Examples of qualitative evidence for the effectiveness of weir removal and related instream modifications (no quantitative evidence found in the reviewed literature)

Reference	Type	Qualitative	Quantitative
Kanehl *et al.* (1997)	Active restoration	After dam removal depth varied considerably following flow variations, rocky bottom increased upstream, bank stability increased upstream and decreased downstream, habitat quality index scores increased dramatically. Short-term effects on fish biomass: increase upstream/long-term effects: general increase.	
Bushaw-Newton *et al.* (2002)	Active restoration	Increased sediment transport has led to major changes in channel form in the former impoundment and downstream reaches, leading benthic macroinvertebrate and fish assemblages to shift dramatically from lentic to lotic taxa. No significant upstream–downstream differences in dissolved oxygen, temperature, or most forms of nitrogen (N) and P, were observed either before or after dam removal.	
Hart *et al.* (2002)	Review	The overall objectives of this article are to assess the current understanding of ecological responses to dam removal and to develop a new approach for predicting dam removal outcomes based on stressor–response relationships.	
Pizzuto (2002)	Review	If the impoundment contains relatively little sediment and is significantly wider than equilibrium channels upstream and downstream of the dam, then the primary processes above the dam are likely to be deposition and floodplain construction rather than erosion and incision. Increased sediment supply at the reach scale could destroy alternate bars, pools and riffles, and armoured beds.	
Shafroth *et al.* (2002)	Review	Following dam removal, large areas of former reservoir bottom are exposed upstream and may be colonised by riparian plants. Transport of upstream sediment may lead to a pulse of sediment deposition downstream, which combined with increased flooding, may both stress existing vegetation and create sites for colonisation and establishment of new vegetation.	
Pollard and Reed (2004)	Active restoration	Cobble habitat without silt generally supports higher taxonomic diversity than do silted areas.	

(*continued*)

Table 3 (*continued*)

Reference	Type	Qualitative	Quantitative
Doyle *et al.* (2005)	Review	Changes in channel form affect riparian vegetation, fish, macroinvertebrates, mussels, and nutrient dynamics.	
Thomson *et al.* (2005)	Active restoration	Downstream sedimentation following dam removal can reduce densities of macroinvertebrates and benthic algae and may reduce benthic diversity, but for small dams such impacts may be relatively minor and will usually be temporary; benthic macroinvertebrate density was significantly lower at downstream sites after complete removal than during pre-removal or partial removal stages, but remained relatively constant at upstream sites (ANOVA); benthic macroinvertebrate assemblages were studied using the NMDS ordination method.	
Cheng and Granata (2007)	Active restoration	After weir removal, net sediment deposition occurred downstream of the dam, and net erosion occurred in the reservoir resulting in bed deposition, and scouring in the reservoir accounted for a decrease in the bed slope of 30% and bed material sizes downstream at least 40% finer than pre-removal conditions; bed deposition and scouring in the reservoir accounted for a decrease in the bed slope of 30% and bed material sizes downstream were at least 40% finer than under pre-removal conditions.	
Maloney *et al.* (2008)	Active restoration	Following the breach, relative abundance of Ephemeroptera, Plecoptera and Trichoptera (largely due to hydropsychid caddisflies) increased upstream, probably because the increased flow and particle size in former impoundments favour filter feeding taxa that cling to substrate (e.g. hydropsychid caddisflies).	
Burroughs *et al.* (2009)	Active restoration	Sediment fill incision resulted in a narrower and deeper channel upstream, with higher mean water velocity and somewhat coarser substrates. Downstream deposition resulted in a wider and shallower channel, with little change in substrate size composition. Water velocity also increased downstream because of the increased slope that developed.	
Tszydel *et al.* (2009)	Active restoration	Riparian and land plants developed intensively at the bottom of the Drzewieckie Reservoir immediately after it was emptied. Short-term flow fluctuations usually diminish the quality and quantity of benthos.	

C. What Is the Timescale of Recovery?

Recovery of the longitudinal connectivity after removal of a weir or dam is immediate, as is the effect on water temperature: it will immediately start changing back to natural (free flowing) conditions. In contrast, biological recovery in general requires several years or even decades after removal and is expected to occur once the fine sediments have been transported farther downstream (e.g. Thomson *et al.*, 2005). This effect depends largely on the quantity of sediments accumulated above the barrier, water velocity, the gradient of the riverbed, and also on the specific technique of weir (dam) removal (Bednarek, 2001). According to Bednarek (2001), full recovery may take up to 80 years, but the literature rarely includes *post hoc* monitoring for longer than 5 years. The timescale and trajectory of recovery after weir removal thus remains speculative, in the absence of long-term monitoring data.

D. Examples of Failure and Limiting Factors When Removing Weirs

Many organisms are limited in their recovery by restricted habitat availability and potential barriers within the river channel. A re-establishment of habitat variability requires geomorphological processes similar to pre-damming conditions (Doyle *et al.*, 2005), which may be key for facilitating fish reproduction, which is often limited by a shortage of suitable habitats to complete their life cycle (i.e. habitat for spawning, nursery, foraging). If geomorphological degradation, however, is irreversible, ecological recovery will hardly be possible without the management of natural geomorphological and hydrological processes (e.g. sediment and flow dynamics). The size of the barrier is also critical: Orr *et al.* (2006) concluded that the initially negative effects of the removal of small dams were trivial and short-lived relative to the natural variability of the entire system (see also Thomson *et al.*, 2005).

Appendix A. Conceptual models and related references

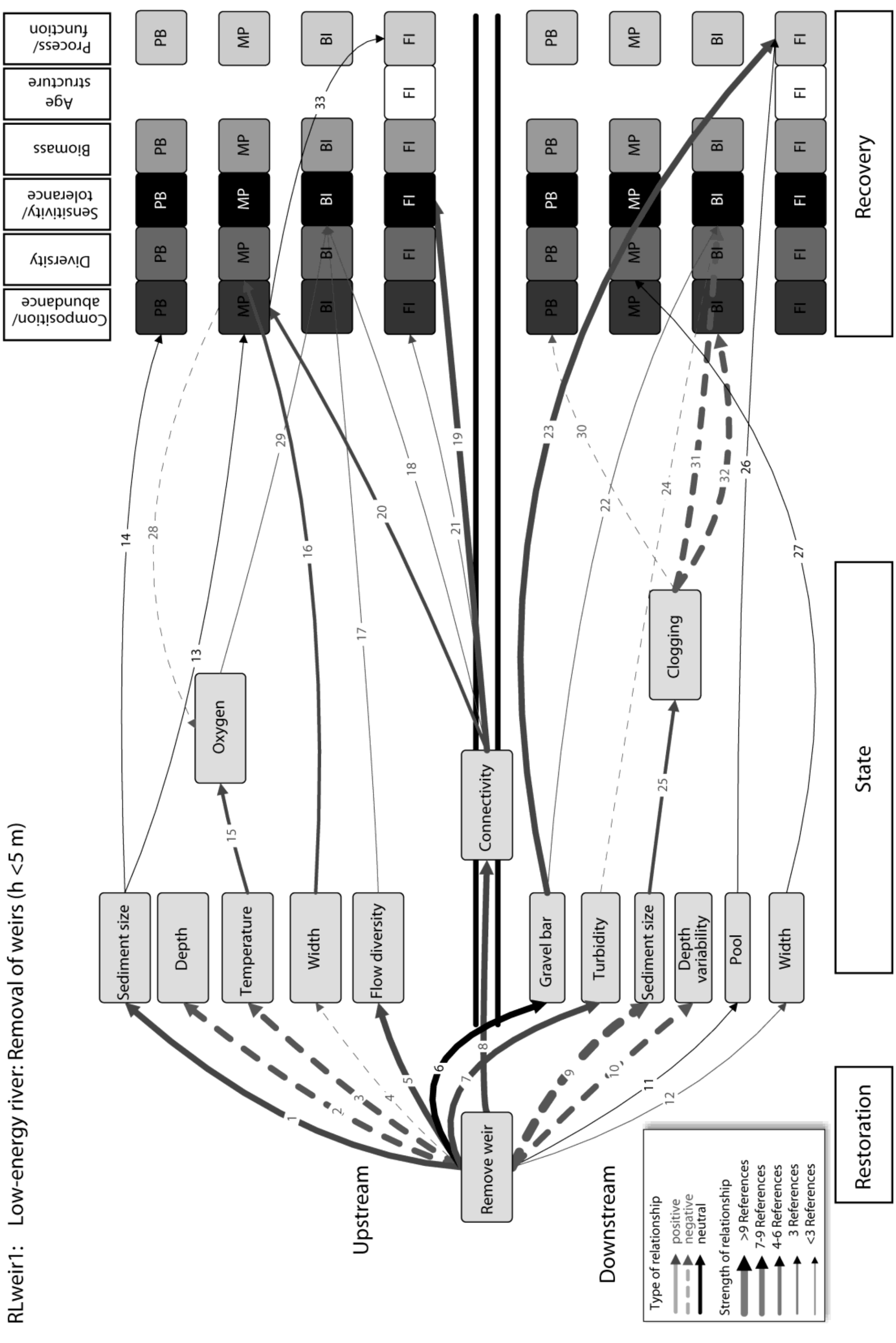


Figure A1. Conceptual models for weir and small dams removal effects on river states and biological communities

Table A1. Matrix of arrow (link) numbers of the conceptual model on weir removal (Figure A1) and restoration studies that refer to the links

A	Model link No.																																	
Serial No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	First author	Year
1																										x								Schlosser	1982
2																										x								Harvey	1991
3																			x		x													Iversen	1993
4			x		x										x																		x	Hill	1993
5	x																																	Kanehl	1997
6								x																										Poff	1997
7													x	x																				Baattrup-Pedersen	1999
8																							x											Bednarek	2001
9		x	x		x					x					x														x		x			Bushaw-Newton	2002
10			x		x	x		x											x	x	x		x											Gregory	2002
11		x		x	x			x	x							x			x															Hart	2002
12																																		Pizzuto	2002
13				x												x				x							x							Shafroth	2002
14			x								x				x							x				x								Stanley	2002
15		x	x				x								x													x						Chaplin	2003
16	x					x	x		x														x		x									Hart	2003
17	x	x				x	x		x	x																								Randle	2003
18						x	x		x	x	x	x								x					x									Rathburn	2003
19	x								x																x						x	x		Pollard	2004
20																																		Doyle	2005
21								x											x															Schmitt	2005
22									x														x							x	x	x		Thomson	2005
23																								x						x	x	x		Orr	2006
24							x		x	x																								Cheng	2007
25													x	x																				Kuhar	2007
26																			x															Leaniz	2008
27																	x	x																Maloney	2008
28		x		x						x	x	x																						Burroughs	2009
29									x		x																							Ahearn	2005
30																																		Ashley	2006
31				x	x																													Evans	2007
32																																		Velinsky	2006
33																			x		x													Stanley	2008
34									x																x					x	x	x		Orr	2008
35	x			x	x				x			x																						Rumschlag	2007
36										x												x									x	x		Tzsydel	2009

Appendix B. References related to Figure A1 and Table A1

Ahearn, D.S. & Dahlgren, R.A. (2005) Sediment and nutrient dynamics following a low-head dam removal at Murphy Creek, California. *Limnology and Oceanography*, **50**(6), 1752-62.

Ashley, J.T.F.,BushawNewton, K., Wilhelm,M.,Boettner, A., Drames,G. & Velinsky, D.J.(2006) The effects of small dam removal on the distribution of sedimentary contaminants.Environmental Monitoring and Assessment, 114(1-3), 287-312.

Baattrup-Pedersen, A. & Riis, T. (1999) Macrophyte diversity and composition in relation to substratum characteristics in regulated and unregulated Danish streams. Freshwater Biology, 42(2), 375-85.

Bednarek, A.T. (2001) Undamming rivers: A review of the ecological impacts of dam removal. Environmental Management, 27, 803-814

Burroughs, B.A., Hayes, D.B., Klomp, K.D., Hansen, J.F. & Mistak, J. (2009) Effects of Stronach Dam removal on fluvial geomorphology in the Pine River, Michigan, United States. Geomorphology, 110(3-4), 96-107.

Bushaw-Newton, K.L., Hart, D.D., Pizzuto, J.E., Thomson, J.R., Egan, J., Ashley, J.T., Johnson, T.E., Horwitz, R.J., Keeley, M., Lawrence, J., Charles, D., Gatenby, C., Kreeger, D.A., Nightengale, T., Thomas, R.L. & Velinsky, D.J. (2002) An integrative approach towards understanding ecological responses to dam removal: The Manatawny Creek Study. Journal of the American Water Resources Association, 38(6), 1581-99

Chaplin, J.J. 2003 , Framework for monitoring and preliminary results after removal of Good Hope Mill Dam. In: W.L. Graf, Editor, Dam Removal Research: Status and Prospects, The H. John Heinz III Center for Science, Economics and the Environment, Washington, DC (2003).

Cheng, F. & Granata, T. (2007) Sediment transport and channel adjustments associated with dam removal: Field observations. Water Resources Research, 43(3).

Doyle, M.W., Stanley, E.H., Orr, C.H., Selle, A.R., Sethi, S.A. & Harbor, J.M. (2005) Stream ecosystem response to small dam removal: Lessons from the Heartland. Geomorphology, 71(1-2), 227-44.

Evans, J.E., Huxley, J.M. & Vincent, R.K. (2007) Upstream channel changes following dam construction and removal using a GIS/remote sensing approach. Journal of the American Water Resources Association, 43(3), 683-97.

Gregory, S., Li, H. & Li, J.(2002) The conceptual basis for ecological responses to dam removal. Bioscience, 52(8), 713-23.

Hart, D.D., Johnson, T.E., Bushaw-Newton, K.L., Horwitz, R.J., Bednarek, A.T., Charles, D.F., Kreeger, D.A. & Velinsky, D.J. (2002) Dam removal: Challenges and opportunities for ecological research and river restoration. Bioscience, 52(8), 669-81.

Hart, 2003 D.D. Hart, T.E. Johnson, K.L. Bushaw-Newton, R.J. Horwitz and J.E. Pizzuto, Ecological effects of dam removal: an integrative case study and risk assessment framework for prediction. In: W.L. Graf, Editor, Dam Removal Research: Status and Prospects, The H. John Heinz III Center for Science, Economics and the Environment, Washington, D.C. (2003).

Harvey, B.C. & Stewart, A.J. (1991) Fish Size and Habitat Depth Relationships in Headwater Streams. Oecologia, 87(3), 336-42.

Hill, M. J., Long, E. A., and Hardin, S. (1994) Effects of dam removal on Dead Lake, Chipola River, Florida. Proceedings of the Annual Conference of Southeastern Association Fish and Wildlife Agencies 48:512–523.

Iversen TM, Kronvang B, Madsen BL, Markmann P, Nielsen MB (1993) Re-establishment of Danish streams: Restoration and maintenance measures. Aquatic Conservation: Marine and Freshwater Ecosystems 3:73–92.

Kanehl, P.D., Lyons, J. & Nelson, J.E. (1997) Changes in the habitat and fish community of the Milwaukee River, Wisconsin, following removal of the Woolen Mills Dam. North American Journal of Fisheries Management, 17(2), 387-400.

Kuhar, U., Gregorc, T., Rencelj, M., Sraj-Krzic, N. & Gaberscik, A. (2007) Distribution of macrophytes and condition of the physical environment of streams flowing through agricultural landscape in north-eastern Slovenia. Limnologica - Ecology and Management of Inland Waters, 37(2), 146-54.

de Leaniz, C.G. (2008) Weir removal in salmonid streams: implications, challenges and practicalities. Hydrobiologia, 609, 83-96.

Maloney, K.O., Dodd, H.R., Butler, S.E. & Wahl, D.H. (2008) Changes in macroinvertebrate and fish assemblages in a medium-sized river following a breach of a low-head dam. Freshwater Biology, 53(5), 1055-68.

Orr, C.H., Rogers, K.L. & Stanley, E.H. (2006) Channel morphology and P uptake following removal of a small dam. Journal of the North American Benthological Society, 25(3), 556-68.

Orr, C.H., Kroiss, S.J., Rogers, K.L. & Stanley, E.H. (2008) Downstream benthic responses to small dam removal in a coldwater stream. River Research and Applications, 24(6), 804-22.

Pizzuto, J. (2002) Effects of dam removal on river form and process. Bioscience, 52(8), 683-91.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E. & Stromberg, J.C. (1997) The Natural Flow Regime. Bioscience, 47, 769-784.

Pollard, A.I. & Reed, T. (2004) Benthic invertebrate assemblage change following dam removal in a Wisconsin stream. Hydrobiologia, 513(1-3), 51-58.

Randle, 2003 T.J. Randle, Dam removal and sediment management Chapter 6. In: W.L. Graf, Editor, Dam Removal Research: Status and Prospects, The Heinz Center for Science, Economics, and the Environment, Washington, D.C. (2003).

Rathburn and Wohl, 2003 S.L. Rathburn and E.E. Wohl, Sedimentation hazards downstream from reservoirs. In: W.L. Graf, Editor, Dam Removal Research: Status and Prospects, The Heinz Center for Science, Economics, and the Environment, Washington, D.C (2003), pp. 105–118.

Rumschlag, J.H. & Peck, J.A. (2007) Short-term sediment and morphologic response of the middle Cuyahoga River to the removal of the Munroe Falls Dam, Summit County, Ohio. Journal of Great Lakes Research, 33(SI2), 142-53.

Schlosser, I.J. (1982) Fish Community Structure and Function Along 2 Habitat Gradients in a Headwater Stream. Ecological Monographs, 52(4), 395-414.

Schmitt, F., Monnier, D., (2005) Impact écologique de l'effacement des barrages dans le Grand-Est. Conseil supérieur de la pêche

Shafroth, P.B., Friedman, J.M., Auble, G.T., Scott, M.L. & Braatne, J.H. (2002) Potential responses of riparian vegetation to dam removal. Bioscience, 52(8), 703-12.

Stanley, E.H., Luebke, M.A., Doyle, M.W. & Marshall, D.W. (2002) Short-term changes in channel form and macro invertebrate communities following low-head dam removal. Journal of the North American Benthological Society, 21(1), 172-87.

Stanley, E.H., Catalano, M.J., Mercado-Silva, N. & Orra, C.H. (2007) Effects of dam removal on brook trout in a Wisconsin stream. River Research and Applications, 23(7), 792-98.

Thomson, J.R., Hart, D.D., Charles, D.F., Nightengale, T.L. & Winter, D.M. (2005) Effects of removal of a small dam on downstream macroinvertebrate and algal assemblages in a Pennsylvania stream. Journal of the North American Benthological Society, 24(1), 192-207.

Tszydel, M., Grzybkowska, M. & Kruk, A. (2009) Influence of dam removal on trichopteran assemblages in the lowland Drzewiczka River, Poland. Hydrobiologia, 630(1), 75-89.

Velinsky, D.J., Bushaw-Newton, K.L., Kreeger, D.A. & Johnson, T.E. (2006) Effects of small dam removal on stream chemistry in southeastern Pennsylvania. Journal of the North American Benthological Society, 25(3), 569-82.

Abstract

Sensitive biological measures of ecosystem quality are needed to assess, maintain or restore the ecological conditions of rivers. Since our understanding of these complex systems is imperfect, river management requires recognizing variability and uncertainty of bio-assessment for decision-making. Based on the analysis of national data sets (~ 1654 sites), the main goals of this work were (1) to test some of the assumptions that shape bio-indicators and (2) address the temporal variability and the uncertainty associated to prediction of reference conditions.

(1) This thesis highlights (i) the predominant role of physiographic factors in shaping biological communities in comparison to human pressures (defined at catchment, riparian corridor and reach scales), (ii) the differences in the responses of biological indicators to the different types of human pressures (water quality, hydrological, morphological degradations) and (iii) more generally, the greatest biological impacts of water quality alterations and impoundments.

(2) A Bayesian method was developed to estimate the uncertainty associated with reference condition predictions of a fish-based bio-indicator (IPR+). IPR+ predictive uncertainty was site-dependent but showed no clear trend related to the environmental gradient. By comparison, IPR+ temporal variability was lower and sensitive to an increase of human pressure intensity.

This work confirmed the advantages of multi-metric indexes based on functional metrics in comparison to compositional metrics. The different sensitivities of macrophytes, fish, diatoms and macroinvertebrates to human pressures emphasize their complementarity in assessing river ecosystems. Nevertheless, future research is needed to better understand the effects of interactions between pressures and between pressures and the environment.

Key-words: Bio-indication ▪ Rivers ▪ Fish ▪ Macroinvertebrates ▪ Benthic diatoms ▪ Macrophytes ▪ Uncertainty ▪ Bayesian modeling ▪ Inter-annual variability ▪ Environmental variability ▪ Reference condition ▪ Water Framework Directive.

Résumé

Evaluer, maintenir et restaurer les conditions écologiques des rivières nécessitent des mesures du fonctionnement de leurs écosystèmes. De par leur complexité, notre compréhension de ces systèmes est imparfaite. La prise en compte des incertitudes et variabilités liées à leur évaluation est donc indispensable à la prise de décision des gestionnaires. En analysant des données nationales (~ 1654 sites), les objectifs principaux de cette thèse étaient de (1) tester certaines hypothèses intrinsèques aux bio-indicateurs et (2) d'étudier les incertitudes de l'évaluation écologique associées à la variabilité temporelle des bio-indicateurs et à la prédiction des conditions de référence.

(1) Ce travail met en évidence (i) le rôle prépondérant des facteurs environnementaux naturels dans la structuration des communautés aquatiques en comparaison des facteurs anthropiques (définis à l'échelle du bassin versant, du corridor ripariens et du tronçon), (ii) les réponses contrastées des communautés aquatiques aux pressions humaines (dégradations hydro-morphologiques et de la qualité de l'eau) et (iii) plus généralement, les forts impacts des barrages et de l'altération de la qualité de l'eau sur les communautés aquatiques.

(2) Une méthode Bayésienne a été développée pour estimer les incertitudes liées à la prédiction des conditions de référence d'un indice piscicole (IPR+). Les incertitudes prédictives de l'IPR+ dépendent du site considéré mais aucune tendance claire n'a été observée. Par comparaison, la variabilité temporelle de l'IPR+ est plus faible et semble augmenter avec l'intensité des perturbations anthropiques.

Les résultats de ce travail confirment l'avantage d'indices multi-métriques basés sur des traits fonctionnels par rapport à ceux relatifs à la composition taxonomique. Les sensibilités différentes des macrophytes, poissons, diatomées et macro-invertébrés aux pressions humaines soulignent leur complémentarité pour l'évaluation des écosystèmes fluviaux. Néanmoins, de futures recherches sont nécessaires à une meilleure compréhension des effets d'interactions entre types de pressions et entre pressions humaines et environnement.

Mots clés : Bio-indication ▪ Rivières ▪ Poissons ▪ Macro-invertébrés ▪ Diatomées benthiques ▪ Macrophytes ▪ Incertitudes ▪ Model bayésien ▪ Variabilité interannuelle ▪ Variabilité environnementale ▪ Condition de référence ▪ Directive Cadre sur l'Eau.