

Un système d'information pour le suivi et l'évaluation de la qualité des cours d'eau

Corinne Grac, Agnès Braud, Florence Le Ber, Michèle Trémolières

▶ **To cite this version:**

Corinne Grac, Agnès Braud, Florence Le Ber, Michèle Trémolières. Un système d'information pour le suivi et l'évaluation de la qualité des cours d'eau. Actes du 3ème atelier " Systèmes d'Information et de Décision pour l'Environnement" - SIDE 2010 - Congrès Inforsid, 2010, Marseille, France. pp.12–21. hal-00498545

HAL Id: hal-00498545
https://hal.science/hal-00498545

Submitted on 7 Jul 2010

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

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

Un système d'information pour le suivi et l'évaluation de la qualité des cours d'eau

C. Grac*, A. Braud**, F. Le Ber*,***, M. Trémolières*

*LHyGeS UMR 7517 - ENGEES, UdS, CNRS, F 67000 Strasbourg
{corinne.grac, florence.leber}@engees.unistra.fr, tremolie@unistra.fr
**LSIIT UMR 7005, UdS, CNRS, F 67400 Illkirch
agnes.braud@unistra.fr
***LORIA UMR 7503, F 54500 Vandœuvre-lès-Nancy

Résumé. La directive cadre européenne sur l'eau (2000) impose la mise au point de nouveaux outils pour l'évaluation et le suivi de la qualité des masses d'eau. Dans ce but, nous avons réalisé depuis 2005 divers prélèvements sur un ensemble de stations (en rivières) de la plaine d'Alsace. Une base de données a été conçue pour organiser et partager les informations collectées : informations relatives aux espèces présentes dans les cours d'eau alsaciens et résultats des prélèvements (physiques, chimiques et biologiques) effectués sur les stations. Pour simplifier et enrichir l'analyse de ces informations nous avons développé un ensemble d'outils comprenant : l'interrogation de la base, la visualisation sur carte des stations respectant des critères donnés, un classifieur traitant les caractéristiques des stations. Notre objectif à terme est de constituer un système d'information utilisable à l'échelle du bassin Rhin-Meuse.

Abstract. The European Water Framework Directive (2000) requires the development of new tools for monitoring and assessing the quality of waterbodies. Following this aim, we collected since 2005 various data from selected sites (of streams) in the Alsace Plain. A database was built to organise and share the whole information : information about the species living in alsacian streams and (physical, chemical and biological) data collected on the sites. Besides we developed tools to facilitate and enrich the analysis of this information : a query tool, a map-based visualisation tool, and a classification tool based on site characteristics. Our further aim is to develop an information system that could be used in the Rhin-Meuse watershed.

Mots-clefs. état biologique des masses d'eau, rivières, indices biologiques, système d'information

Keywords. biological quality of waterbodies, streams, biological indices, information system

Introduction

La qualité des eaux de surface est un problème majeur en Europe, comme l'a souligné la Directive Cadre Européenne sur l'Eau (DCE), datant de l'année 2000. L'évaluation de la qualité de l'eau sur les seuls critères physico-chimiques est apparue comme insuffisante depuis les années 70 et l'usage complémentaire d'outils biologiques, tel que le premier indice français, l'IBGN (Indice Biologique Global Normalisé), basé sur sur les invertébrés, s'est généralisé depuis 1992 [1]. Depuis 2000, quatre autres indices ont été normalisés en France, mais leur utilisation conjointe afin d'évaluer l'état d'un écosystème dans son ensemble n'a pas été immédiate [3] et elle n'a débuté que depuis peu.

Notre projet a pour but de proposer un tel outil global d'évaluation, au moyen d'un système d'information permettant de gérer et d'analyser les différentes données concernant les systèmes aquatiques de la plaine d'Alsace. Pour cela, nous avons d'abord développé une base de données regroupant les données existantes, collectées sur environ 700 stations de l'hydroéco-région de la plaine d'Alsace [18] depuis 20 ans pour les plus anciennes : il s'agit de données physiques, physico-chimiques, floristiques et faunistiques. La base inclut également les informations utiles à l'analyse et la synthèse de ces données, en particulier les caractéristiques des taxons (traits biologiques et écologiques) ainsi que les seuils de qualité (physique, chimique et biologique) et les valeurs des différents indices biologiques français calculés sur les stations. Par la suite, la base a été dotée d'une interface permettant d'accompagner les biologistes dans leurs analyses, au travers d'outils de recherche et de recoupement des informations stockées dans la base. Cette interface autorise différentes vues et en particulier un accès cartographique de l'information. Finalement nous développons un classifieur permettant de comparer les caractéristiques des stations et d'en donner une évaluation globale. La base de données originelle est ainsi complétée d'un ensemble d'outils cohérents qui en font un véritable système d'information pour l'évaluation et le suivi de la qualité des cours d'eau.

Cet article présente les différents aspects du système d'information ainsi développé. La première partie s'attache à la description de la base de données, la seconde à la description des interfaces avancées et la dernière à la description du classifieur en cours de développement. Nous discutons de notre approche puis concluons sur l'intérêt et les perspectives de ce projet.

1 Structure et contenu de la base de données

La base contient différents types de données sur les stations de cours d'eau : des données environnementales telles que le débit et le temps (climat) au moment du prélèvement ; des données physiques, concernant l'état hydromorpho-

logique du cours d'eau ; des données chimiques, telles que les taux de nitrates, phosphates, matières organiques présents dans l'eau ; des données floristiques ; des données faunistiques. Les données floristiques recouvrent les diatomées (algues) et les macrophytes (ou hydrophytes). Les données faunistiques concernent les invertébrés, les oligochètes benthiques et les poissons. Une partie des données a été collectée par le laboratoire LHyGeS (et anciennement le CEVH), une autre partie provient d'organismes publics tels que l'Office National de l'Eau et des Milieux Aquatiques (ONEMA) et l'Agence de l'Eau Rhin Meuse (AERM). Les plus anciennes données concernent essentiellement les aspects physico-chimiques et les macrophytes [14, 13]. Les données récentes, collectées dans le cadre du projet INDICES (2005-2009) [9] concernent tous les compartiments biologiques pour une quarantaine de stations, choisies parmi les sept types de cours d'eau présents en plaine d'Alsace. Les méthodes d'échantillonnage utilisées pour les relevés floristiques et faunistiques sont les méthodes normalisées des indices biologiques, modifiées selon les recommandations du programme européen de recherches AQEM[1] ; pour les invertébrés, nous avons suivi le protocole établi dans [16].

La base de données développée [8] obéit au format national SANDRE[2] pour les données aquatiques. Elle contient 38 tables. Les principales tables concernent la description des stations et des données physico-chimiques et hydrobiologiques : stations échantillonnées, dates d'échantillonnage, conditions environnementales, méthodes d'échantillonnage, résultats chimiques, résultats biologiques (voir figure 1).

Les autres tables contiennent des informations sur les différents paramètres, en particulier les paramètres chimiques et biologiques. Par exemple, chaque taxon floristique ou faunistique est représenté par sa nomenclature taxonomique dans une table Taxon. Cette table provient du SANDRE (environ 2000 enregistrements à l'époque de la création de la base) et a été complétée (environ 4500 enregistrements). Trois tables reliées à la table Taxon détaillent les caractéristiques de chaque taxon. Les informations concernant les traits biologiques et écologiques des taxons présents en Alsace ont également été enregistrées dans des tables spécifiques. Ces informations ont été collectées dans la littérature [19, 15, 12, 17] et adaptées ou complétées pour le contexte local.

La base a été implantée sous MySQL. Elle a été alimentée régulièrement au cours du projet INDICES avec des données existantes et suite aux prélèvements effectués et analyses faites en laboratoire. À l'heure actuelle, les données concernant quasiment tous les prélèvements effectués au cours du projet sont saisies et tous les paramètres utiles à l'évaluation de la qualité de l'eau sont renseignés, permettant ainsi une analyse complète sur un ensemble de stations représentatif de la plaine d'Alsace.

[1] http ://www.aqem.de/
[2] http ://sandre.eaufrance.fr/

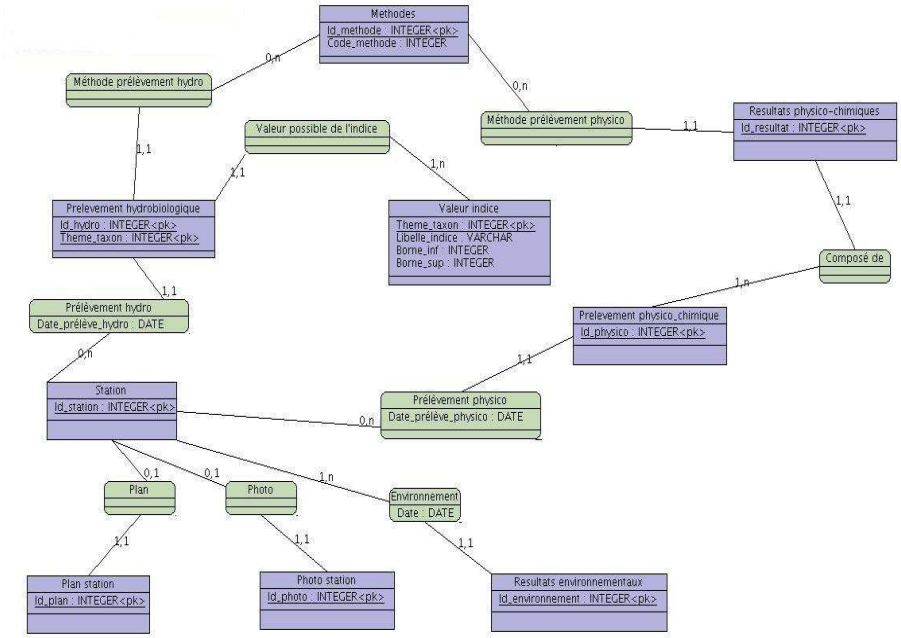


FIG. 1 – MCD partiel de la base de données : les tables décrivant les stations, les méthodes d'échantillonnage et les résultats, avec leurs liens

2 Interfaces avancées

Les données enregistrées dans la base sont utilisées par différentes personnes, enseignants-chercheurs et étudiants pour le moment, mais ingénieurs ou techniciens ultérieurement. L'entrée naturelle, pour ces utilisateurs, est une entrée cartographique, car elle permet de relier le contenu de la base au terrain qu'ils ont enquêté ou qu'ils veulent diagnostiquer. Les données enregistrées dans la base sont également utilisées dans différentes perspectives : recherche de stations ayant les mêmes caractéristiques (mêmes valeurs d'indices biologiques, par exemple), possédant tels ou tels taxons (macrophytes, poissons, ...), échantillonnées à telle ou telle période ; ou bien recherche des stations où se trouvent tels ou tels taxons à différentes périodes, etc. Enfin, les utilisateurs veulent pouvoir extraire les résultats de leurs requêtes sous forme de tableaux sur lesquels ils pourront ensuite faire des analyses.

Ces besoins ont été établis progressivement et nous ont conduites à développer une interface cartographique doublée d'une interface de requête avancée, permettant de sélectionner simultanément un ensemble de stations *via* la carte, puis de rechercher les taxons s'y trouvant. Outre ces informations, le tableau extrait peut contenir des informations synthétiques, calculées à la volée (consultation

des stations, figure 2). Symétriquement, nous développons une vue permettant de sélectionner des taxons et de visualiser puis extraire les stations où ils se trouvent (consultation des taxons). L'ensemble est accessible via un site web, rendu nécessaire par la dispersion physique des utilisateurs. Le choix d'une interface *ad hoc* plutôt que d'un couplage de la base avec un système d'information géographique s'est fondé sur différentes raisons dont nous discutons plus loin.



FIG. 2 – Vue de l'interface : consultation des stations

3 Principes du classifieur

Un des objectifs du projet INDICES, et donc de la constitution du système d'information, est de comparer et combiner les valeurs d'indices biologiques. Pour un ensemble de 40 stations, nous possédons en effet les informations nécessaires pour calculer les cinq indices biologiques normalisés français, à savoir l'IBD (établi sur les diatomées), l'IBGN (invertébrés), l'IBMR (macrophytes), l'IOBS (oligochètes) et l'IPR (poissons). Ces valeurs d'indices sont ensuite transformées en cinq classes de qualité (cf. figure 2, en bas de l'interface).

L'idée du classifieur, issue de réflexions sur le système d'harmonisation proposé dans [10], est d'évaluer la qualité d'une station en l'associant automatiquement à un "profil caractérisé". Ce profil correspond à un ensemble de stations connues, stockées dans le système d'informations, et partageant des caractéristiques communes, à partir desquelles on peut définir un profil. Par

exemple, les stations possédant les classes de qualité suivantes (IBGN=[1,2], IOBS=3, IBMR=3, IBD=[4,5]) sont associées au profil suivant : *"début de dégradation des sédiments, forte dégradation physico-chimique au moins liée à un niveau trophique moyen, mais hors matière organique, bon potentiel de résilience général et possibilité de résilience sur les sédiments"*. L'intérêt ici pour les biologistes est de disposer d'une évaluation globale des stations, y compris dans les compartiments dont ils ne sont pas spécialistes. De plus, ce classifieur permet d'observer l'évolution temporelle d'une station pour laquelle on dispose d'un suivi pluriannuel.

Pour déterminer de tels ensembles de stations, nous utilisons la technique des treillis de Galois [2]. Un algorithme de construction de treillis de Galois prend en entrée un ensemble d'objets, un ensemble de propriétés, et une relation d'incidence précisant pour chaque couple (objet, propriété) si l'objet possède la propriété. Il fournit en sortie un ensemble hiérarchisé de concepts, c'est-à-dire des groupes d'objets définis par leurs propriétés communes. Il permet également de générer les règles d'association entre ces propriétés.

La démarche de construction du treillis est présentée dans [6]. Les propriétés prises en compte pour définir les groupes ont été dans un premier temps les cinq indices biologiques normalisés, et nous travaillons à l'heure actuelle sur l'intégration des paramètres physico-chimiques. La structure hiérarchique obtenue permettra alors de parcourir efficacement l'ensemble des concepts afin de répondre à des requêtes, comme cela a été fait dans les domaines biologique ou géographique [11, 4]. Nous comptons utiliser une approche similaire pour évaluer une nouvelle station, en déterminant le concept auquel elle appartient ou ceux dont elle se rapproche le plus.

4 Discussion

Les fonctionnalités proposées dans ce système d'information viennent combler un manque d'outils exprimé par les chercheurs en biologie du LHyGeS pour ce qui concerne certaines de leurs tâches.

Un premier problème concerne le stockage et la manipulation des données collectées sur le terrain. Elles sont stockées classiquement dans des feuilles de tableur de grande taille et les recoupements sont effectués à la main, engendrant un travail fastidieux et des risques d'erreurs. À l'inverse, le système de gestion de base de données apporte des garanties en termes de cohérence et de sécurité des données. Le système d'information offre également des fonctionnalités de requêtes avancées qui apportent un mécanisme puissant de recherche et de recoupement de données issues de sources multiples.

Un second problème concerne la création de cartes permettant de visualiser la répartition géographique des valeurs des paramètres mesurés sur les stations. Jusqu'à présent ces cartes sont créées à la main : un tableau contenant les informations à visualiser doit être constitué (étape manuelle) puis enregistré dans un système d'information géographique (SIG) qui affiche alors les données thématiques sur la carte. Grâce à l'interface cartographique du système d'infor-

mation, les cartes sont construites par une simple requête et, de plus, l'ensemble des informations concernant une station ou plusieurs stations sélectionnées est accessible par simple clic sur ces stations.

Pour répondre à ces différents besoins, et en particulier aux besoins cartographiques, il a été tout d'abord envisagé d'utiliser un SIG, couplé à la base de données. Nous avons mené une étude comparative confrontant différents SIG et un développement *ad hoc* [7]. Nous avons considéré uniquement des SIG Open Source, tels que GRASS[3], GeOxygene[4], ou GvSIG[5], afin de pouvoir adapter le SIG choisi à nos besoins. À l'issue de cette étude, nous avons opté pour un développement *ad hoc*, pour les raisons suivantes :

- les données considérées ont une spatialité faible, seules les stations doivent être localisées et ceci par un point, leur extension spatiale n'étant pas précisée ;
- les besoins en fonctionnalités cartographiques se limitent à l'heure actuelle au positionnement de stations sur une carte, et éventuellement plus tard à la détermination d'un voisinage ;
- en revanche de nombreuses fonctionnalités sont nécessaires pour aider les biologistes dans le recoupement d'informations issues de plusieurs sources ;
- la visualisation sur carte est une fonctionnalité supplémentaire offerte aux biologistes pour faciliter la recherche des informations liées aux stations, sachant qu'il est plus facile de les identifier visuellement, par zooms successifs, que par un code.

Il apparaît ainsi que le recouvrement entre les nombreuses fonctionnalités offertes par les SIG et nos besoins spécifiques est assez faible. De plus, utiliser un SIG impose certaines contraintes alors qu'un développement *ad hoc* permet une plus grande souplesse pour obtenir les résultats attendus. En l'occurrence, GvSIG était l'option la plus appropriée, mais il n'y avait pas de version disponible stable à l'époque. Une étude menée aujourd'hui nous conduirait peut-être à des conclusions différentes.

Finalement, et ceci justifie largement notre choix, faire un développement *ad hoc* nous a permis d'adapter en continu l'interface aux biologistes qui en sont destinataires. L'outil est, à leurs yeux, simple, intuitif, et ne nécessite pas de formation. Ceci est d au fait qu'il est conçu en collaboration directe avec les biologistes (chercheurs et étudiants) qui font leurs propres propositions, testent et valident les différents affichages et les fonctionnalités.

5 Conclusions et perspectives

Le système d'information présenté dans cet article a été conçu dans le cadre du projet INDICES et dans l'objectif d'intégrer tous les éléments nécessaires à l'étude d'outils d'évaluation globale de la qualité des cours d'eau tels que demandés par la DCE. Il rassemble pour cela une grande variété d'informations

[3]http ://grass.itc.it/
[4]http ://oxygene-project.sourceforge.net/
[5]http ://www.gvsig.gva.es/

ayant trait à l'évaluation des cours d'eau de la plaine d'Alsace. Les données stockées sont issues de bases de données nationales faisant référence (pour les taxons par exemple), issues d'une synthèse bibliographique réalisées par les biologistes du LHyGeS (traits biologiques), ou résultats d'analyses faites sur des prélèvements. Ces prélèvements ont en particulier été effectués sur un échantillon de stations sélectionnées en plaine d'Alsace pour représenter les différents types de stations identifiés par la DCE.

Pour permettre une exploitation efficace de ce système d'information par les biologistes et les aider à valoriser et recouper l'ensemble des informations, nous avons conçu différents outils d'interrogation simple ou avancée, de visualisation sur carte, de classification automatique de stations. Ces outils sont intégrés dans un site web. L'ensemble est simple d'utilisation et accessible depuis n'importe quel poste sans nécessité d'installer un logiciel autre qu'un navigateur.

Les agences de l'eau ont développé des bases de données où sont recensées les informations sur les nombreuses stations qu'elles surveillent. Certaines offrent une visualisation cartographique, comme l'Agence de l'Eau Loire-Bretagne, avec l'interface Osur Web[6]. Toutefois, même si elles recouvrent des zones géographiques étendues, les informations disponibles dans ces bases sont très limitées, par exemple on ne dispose généralement pas des relevés taxonomiques établis sur les stations.

Sur cette même thématique, l'ONEMA a commandé le projet SEEE-cours d'eau (Système d'Évaluation de l'État de l'Eau). Il s'agit de rassembler des banques de données afin de permettre la mise au point puis l'intégration d'outils d'évaluation. Nos objectifs sont similaires, les principales différences viennent des dimensions des deux projets : le SEEE a un cadre national tandis que nous travaillons sur la plaine d'Alsace ; de plus, le SEEE a pour vocation de réaliser l'évaluation DCE des cours d'eau, autrement dit, en se fondant sur la plus mauvaise valeur, alors que nous considérons tous les indices ensemble.

Par ailleurs, le projet INDICES ne se limite pas aux seules données nécessaires au calcul des indices biologiques ou physico-chimiques, mais vise à intégrer des informations supplémentaires, tels que les traits biologiques et écologiques des taxons, pour construire un outil d'évaluation de la qualité biologique globale des cours d'eau [5]. Finalement nous mettons en œuvre des outils innovants, comme les treillis de Galois, afin d'explorer au mieux les données collectées. L'approche fondée sur le classifieur sera prochainement finalisée et testée sur des données équivalentes recueillies en Bretagne. Ceci pourrait conduire à confronter des expertises différenciées, les espèces considérées pour construire les indices pouvant être plus ou moins présentes dans les différentes régions.

Remerciements. Les auteurs remercient l'Agence de l'Eau Rhin-Meuse et l'ONEMA pour leur soutien à ce projet.

[6] http ://carto.eau-loire-bretagne.fr/osur/

Références

[1] AFNOR. Qualité de l'eau : détermination de l'indice biologique global normalisé (ibgn), 1992. NF T90-350.

[2] M. Barbut et B. Monjardet. *Ordre et classification – Algèbre et combinatoire*. Hachette, 1970.

[3] M.-F. Bazerques. Directive-cadre sur l'eau : le bon état écologique des eaux douces de surface : sa définition, son évaluation. Communication au Ministère de l'Écologie et du Développement Durable, Paris, 2004.

[4] O. Bedel, S. Ferré, O. Ridoux, et E. Quesseveur. GEOLIS : a logical information system for geographical data. *Revue Internationale de Géomatique*, 17 :371–390, 2007.

[5] A. Bertaux, F. Le Ber, A. Braud, et M. Trémolières. Identifying ecological traits : a concrete fca-based approach. In *7th International Conference on Formal Concept Analysis, ICFCA 2009, Darmstadt*, LNAI 5548, pages 224–236. Springer-Verlag, 2009.

[6] A. Braud, C. Grac, S. Pristavu, E. Dor, et F. Le Ber. Une démarche fondée sur les treillis de Galois pour l'aide à la qualification de l'état des milieux aquatiques. In *Actes du 2ème Atelier Systèmes d'Information et de Décision pour l'Environnement - SIDE 2009, Toulouse*, pages 94–105, 2009.

[7] M. Buleandra. Visualisation de données sur la qualité des cours d'eau en Alsace. Mémoire de stage Erasmus, LSIIT, Illkirch, France et Université Dunarea de Jos, Galati, Roumanie, 2007.

[8] J.-L. Ehrhard. Mise en œuvre d'un système de comparaison des réponses des indices biologiques sur les cours d'eau de la plaine d'alsace. Mémoire de diplôme d'ingénieur CNAM en informatique, Strasbourg, 2005. CEVH.

[9] C. Grac, F. Le Ber, A. Herrmann, et M. Trémolières. Programme de recherche-développement *Indices* – rapport d'avancement scientifique de la deuxième année (2008). Contrat pluriannuel 1463 de l'Agence de l'Eau Rhin-Meuse, CEVH, 2009.

[10] M. Lafont, S. Vigneron, et A. Fournier. Evaluation de l'effet des rejets polluants sur les milieux aquatiques situés dans des environnements imperméabilisés : Proposition d'une approche intégrée. Rapport 01-0784, Cemagref, 2001.

[11] N. Messai, M.-D. Devignes, A. Napoli, et M. Smaïl Tabbone. Correction et complétude d'un algorithme de recherche d'information par treillis de concepts. In *Classification : points de vue croisées*, Revue des Nouvelles Technologies de l'Information (RNTI), pages 147–158. Cépaduès Éditions, 2008.

[12] H. Tachet, P. Richoux, M. Bournaud, et P. Usseglio-Polatera. *Invertébrés d'eau douce : Systématique, biologie, écologie*. CNRS Éditions, 2000. 588 pages.

[13] M. Trémolières. Fiches descriptives des habitats aquatiques. In *Référentiel des habitats reconnus d'intért communautaire de la bande rhénane*, pages 73–110. Conservatoire des Sites Alsaciens et Office National de Forts, 2004. Programme LIFE Rhin vivant.

[14] M. Trémolières, R. Carbierner, A. Orstcheit, et J.-P. Klein. Changes in aquatic vegetation in rhine floodplain streams in alsace in relation to disturbance. *Journal of Vegetation Science*, 5 :169–178, 1994.

[15] P. Usseglio-Polatera et J.-N. Beisel. Longitudinal changes in macroinvertebrate assemblages in the meuse river : anthropogenic effects versus natural change. *River Res. Applic.*, 18 :197–211, 2002.

[16] P. Usseglio-Polatera et J.-G. Wasson. Protocole de prélèvement et de traitement des échantillons des macro-invertébrés benthiques sur les sites de référence "cours d'eau". Université de Metz et Cemagref de Lyon, 2004. 7 pages.

[17] H. Van Dam, A. Mertens, et J. Sinkeldam. A coded checklist and ecological indicator values of freshwater diatoms from the netherlands. *Aquatic Ecology*, 28(1) :117–133, 1994.

[18] J.G. Wasson, A. Chandesris, H. Pella, et L. Blanc. Les hydroécorégions de France métropolitaine - approche régionale de la typologie des eaux courantes et éléments pour la définition des peuplements de référence d'invertébrés. Rapport 02-0413, Cemagref, 2002. 190 pages.

[19] N.J. Willby, V.J. Abernethy, et B.O.L. Demars. Attribute-based classification of european hydrophytes and its relationship to habitat utilization. *Freshwater Biology*, 43 :43–74, 2000.