



HAL
open science

**Étude sur le protocole d'échantillonnage de l'IBGA.
Comparaison des méthodologies utilisées pour le
protocole d'échantillonnage des invertébrés en grand
cours d'eau (IBGA) en vue d'une meilleure
représentativité des la faune locale**

M. Ferreol, J.G. Wasson, Virginie Archambault, P. Usseglio Polatera

► **To cite this version:**

M. Ferreol, J.G. Wasson, Virginie Archambault, P. Usseglio Polatera. Étude sur le protocole d'échantillonnage de l'IBGA. Comparaison des méthodologies utilisées pour le protocole d'échantillonnage des invertébrés en grand cours d'eau (IBGA) en vue d'une meilleure représentativité des la faune locale. irstea. 2010, pp.28. hal-02599632

HAL Id: hal-02599632

<https://hal.inrae.fr/hal-02599632>

Submitted on 16 May 2020

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

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

Convention de partenariat ONEMA-Cemagref 2008

Domaine :

Action :

Étude sur le protocole d'échantillonnage de l'I.B.G.A.

Comparaison des méthodologies utilisées pour le protocole d'échantillonnage des invertébrés en grand cours d'eau (IBGA) en vue d'une meilleure représentativité des la faune locale.

Martial Ferréol, Jean-Gabriel Wasson, Virginie Archaimbault,
Philippe Usseglio-Polatera

CEMAGREF
Département Gestion des Milieux Aquatiques
**Unité de recherche Biologie des Ecosystèmes
Aquatiques**
Laboratoire d'Hydroécologie Quantitative
Groupement de Lyon
3bis Quai Chauveau CP 220
69336 LYON CEDEX 09

2010

Résumé

L'évaluation écologique des grands cours d'eau fait actuellement partie des grands chantiers de la mise en œuvre de la Directive Cadre Européenne sur l'eau (DCE). Dans cette perspective, et du fait de leur grand potentiel bio-indicateur, les macro-invertébrés benthiques font partie des éléments de qualité biologique à suivre. Mais la nature de ce type de masse d'eau induit des difficultés au niveau du protocole d'échantillonnage.

Historiquement l'évaluation de la qualité écologique des grands cours d'eau, basée sur les invertébrés, était estimée à partir de l'IBGA (Indice Biologique Global Adapté). Son principe était simple et dérivait de l'IBGN (Indice Biologique Global Normalisé). La faune des rives était échantillonnée en 8 réplicats à l'aide de filet Surber ou troubleau (méthode "sur Berges") et la faune des habitats profonds était échantillonnée à la fois par les substrats artificiels (2 réplicats) et en effectuant des dragages (3 réplicats). Cependant, comme l'IBGN, cette méthodologie ne correspondait pas aux exigences de la DCE (non prise en compte des peuplements de référence, de l'abondance, des ratios entre taxons sensibles et tolérants) et n'était pas comparable aux méthodes utilisées dans les autres pays européens. Il devenait donc urgent de rendre cette méthode compatible d'une part à la DCE et d'autre part aux méthodes appliquées pour les petits cours d'eau afin d'assurer une continuité écologique et un maximum d'homogénéité dans les protocoles d'échantillonnage le long des gradients amont-aval.

Ce travail présente les démarches analytiques qui ont conduit à l'élaboration du choix des outils d'échantillonnage à combiner dans ce nouveau protocole. Il se base sur de nombreux résultats issus des données historiques disponibles de l'IBGA pour plusieurs grands cours d'eau (l'Ain, l'Aisne, le Doubs, l'Isère, la Loire, la Marne, la Meuse, le Rhône, la Saône, la Seille et la Seine) et a pour objectif de comparer l'apport de chacune des méthodes d'échantillonnage pour une meilleure appréhension de la faune totale en macro-invertébrés.

Il s'agit en fait d'étudier la variabilité inter-méthode et non de discuter des méthodes en elles-mêmes. Les analyses portent directement sur le contenu des listes taxonomiques dont les individus ont été identifiés au niveau de la famille. Le principe des comparaisons repose sur le croisement des listes faunistiques issues des trois méthodes d'échantillonnage avec notamment la comparaison des richesses taxonomiques, la comparaison par ordination des populations entre sites d'un même cours d'eau ou encore les estimations des taux d'erreur auxquels on s'expose si on ne pratique pas une des 3 méthodes d'échantillonnage.

Si la méthode dite "sur Berges" (8 prélèvements au filet Surber) est qualitativement et quantitativement toujours la plus significative, chacune a son importance. La méthode par Dragage est une réelle plus value car elle apporte des taxons bio-indicateurs qui ne seront pas forcément prélevés sur les berges. Par contre, la méthode par Substrat Artificiel permet plutôt de palier des éventuels manques à la méthode dite "sur Berges", et pouvant concerner des taxons hautement bio-indicateurs. Cette constatation globale est moins évidente à l'échelle d'une seule rivière où on a remarqué des nuances quant à l'apport faunistique des listes Drague et Substrat Artificiel.

La prévalence de chacune de ces méthodes doit donc être replacée dans son contexte environnemental afin d'obtenir la meilleure image possible d'un grand cours d'eau d'une part et avec le plus faible coût en effort d'échantillonnage d'autre part.

Table des matières

1	INTRODUCTION.....	6
2	DONNEES DISPONIBLES	6
2.1	BASSINS ET RIVIERES CONCERNEES	6
2.2	METHODE D'ECHANTILLONNAGE	8
2.3	DONNEES TAXONOMIQUES	8
2.4	APPROCHE COMPARATIVE	8
3	ANALYSES COMPARATIVES	9
3.1	APPORTS QUANTITATIFS EN TAXONS (RICHESS) SELON LA TECHNIQUE D'ECHANTILLONNAGE	9
3.2	COMPARAISON DES DISTRIBUTIONS TAXONOMIQUES	13
3.3	ÉTUDE DES OCCURRENCES EN TAXONS PAR TECHNIQUES D'ECHANTILLONNAGE	15
3.3.1	<i>Apport sur la liste B de la prise en compte d'une seconde technique d'échantillonnage</i>	<i>15</i>
3.3.2	<i>Apport de la prise en compte d'une troisième technique d'échantillonnage</i>	<i>16</i>
3.3.3	<i>Interprétations taxonomiques</i>	<i>17</i>
3.3.4	<i>Taxons bio-indicateurs.....</i>	<i>20</i>
4	DISCUSSION	23
4.1	COHERENCE DES RESULTATS	23
4.2	DIFFERENCES ENTRE LES METHODES D'ECHANTILLONNAGE	23
5	CONCLUSION.....	24
ANNEXE A.	RAPPELS DE PROBABILITES.....	25
ANNEXE B.	DEMONSTRATION DE LA SIMPLIFICATION DE LA FORMULE DE PROBABILITES.....	27
ANNEXE C.	PRINCIPE DE L'ESTIMATION PAR BOOTSTRAPS DE L'ERREUR STANDARD FAITE SUR L'ESTIMATION DES TAUX D'ERREUR.....	28

Table des figures

FIGURE 1 : CARTE DES COURS D'EAU RETENUS POUR L'ANALYSE DU PROTOCOLE D'ÉCHANTILLONNAGE DE L'IBGA.....	7
FIGURE 2 : SCHEMA REGROUPANT LES DIFFERENTES POSSIBILITES D'OCCURRENCE DES TAXONS PAR SITE PARMIS LES DIFFERENTES METHODES D'ÉCHANTILLONNAGE. LA LOGIQUE EST LA MEME QUE CELLE DU CERCLE DES COULEURS PRIMAIRES ICI REPRIS. LES COULEURS PRIMAIRES CORRESPONDENT AUX CAS OU UN TAXON EST EXCLUSIVEMENT TROUVE DANS UNE SEULE LISTE D'UN SITE. LES COULEURS SECONDAIRES CORRESPONDENT AUX CAS OU UN TAXON DONNE EST EXCLUSIFS A DEUX METHODES D'ÉCHANTILLONNAGE. ENFIN, LE BLANC ET LE NOIR (EN GRIS SUR LE SCHEMA PAR COMMODITE GRAPHIQUE) CORRESPONDENT RESPECTIVEMENT A LA SITUATION OU LE TAXON EST VU DANS LES TROIS LISTES OU AU CONTRAIRE DANS AUCUNE	9
FIGURE 3 : DISTRIBUTIONS DES RICHESSES EN FAMILLE DE MACRO-INVERTEBRES PAR RIVIERE ET SELON LES METHODES D'ÉCHANTILLONNAGE PRISES EN COMPTE.....	12
FIGURE 4 : SCORES DES LISTES DES AFC EFFECTUEES POUR LES RIVIERES DU TYPE 1. LES SCORES SONT REGROUPES PAR METHODE D'ÉCHANTILLONNAGE. LES VALEURS PROPRES DES AFC SONT INDIQUEES EN PROPORTION.....	13
FIGURE 5 : SCORES DES LISTES DES AFC EFFECTUEES POUR LES RIVIERES DU TYPE 2 (DRAGUE). LES SCORES SONT REGROUPES PAR METHODE D'ÉCHANTILLONNAGE. LES VALEURS PROPRES DES AFC SONT INDIQUEES EN PROPORTION DE L'INERTIE TOTALE	14
FIGURE 6 : SCORES DES LISTES DES AFC EFFECTUEES POUR LES RIVIERES DU TYPE 2 (SUBSTRAT ARTIFICIEL). LES SCORES SONT REGROUPES PAR METHODE D'ÉCHANTILLONNAGE. LES VALEURS PROPRES DES AFC SONT INDIQUEES EN PROPORTION DE L'INERTIE TOTALE	14
FIGURE 7 : SCORES DES LISTES DES AFC EFFECTUEES POUR LES RIVIERES DU TYPE 3. LES SCORES SONT REGROUPES PAR METHODE D'ÉCHANTILLONNAGE. LES VALEURS PROPRES DES AFC SONT INDIQUEES EN PROPORTION DE L'INERTIE TOTALE	15
FIGURE 8 : ÉVOLUTION DES VALEURS MOYENNES PAR TAXON DES PROPORTIONS DE SITES BENEFICIANT DE L'AJOUT D'UNE LISTE TAXONOMIQUE ISSUE SOIT DE LA METHODE D (NOIR), SOIT DE LA METHODE SA (ROUGE) PAR RAPPORT A LA LISTE ISSUE DE LA METHODE B. LES VALEURS SONT CLASSEES DANS L'ORDRE DECREISSANT RESPECTIF A CHAQUE METHODE	16
FIGURE 9 : HISTOGRAMMES DES TAUX D'ERREUR CUMULE PAR CATEGORIE DE TAXON ET PAR RIVIERE SELON LA METHODE D'ÉCHANTILLONNAGE : A- BERGES, B- SUBSTRAT ARTIFICIEL ET C- DRAGUE. LES RIVIERES SONT REGROUPEES SELON LES TYPES IDENTIFIES EN 3.2.....	19

Table des tableaux

TABLEAU 1 : TABLE DE CONTINGENCE DES ECHANTILLONS DISPONIBLES PAR BASSIN HYDROGRAPHIQUE.....	6
TABLEAU 2 : TABLE DE CONTINGENCE DES ECHANTILLONS DISPONIBLES PAR COURS D'EAU	7
TABLEAU 3 : TESTS DE COMPARAISON DES MOYENNES APPARIEES SELON LA METHODE D'ECHANTILLONNAGE ENTRE DISTRIBUTIONS DE RICHESSES PAR RIVIERE. LES P-VALUES SONT INDIQUEES ENTRE PARENTHESES. EN JAUNE QUAND LA P-VALUE EST SIGNIFICATIVE (< 0.05), NOTE *. EN ROUGE QUAND LA P-VALUE EST TRES SIGNIFICATIVE (< 0.01), NOTE **	10
TABLEAU 4 : TESTS DE COMPARAISON DES MOYENNES APPARIEES SELON LES COMBINAISONS DE METHODES D'ECHANTILLONNAGE ENTRE DISTRIBUTIONS DE RICHESSES PAR RIVIERE. LES P-VALUES SONT INDIQUEES ENTRE PARENTHESES. EN JAUNE QUAND LA P-VALUE EST SIGNIFICATIVE (< 0.05), NOTE *. EN ROUGE QUAND LA P-VALUE EST TRES SIGNIFICATIVE (< 0.01), NOTE **	10
TABLEAU 5 : TESTS DE COMPARAISON DES MOYENNES APPARIEES SELON LES COMBINAISONS DE METHODES D'ECHANTILLONNAGE ENTRE DISTRIBUTIONS DE RICHESSES PAR RIVIERE. LES P-VALUES SONT INDIQUEES ENTRE PARENTHESES. EN JAUNE QUAND LA P-VALUE EST SIGNIFICATIVE (< 0.05), NOTE *. EN ROUGE QUAND LA P-VALUE EST TRES SIGNIFICATIVE (< 0.01), NOTE **	11
TABLEAU 6 : TESTS DE <i>t</i> DE COMPARAISONS DES MOYENNES DES SCORES GFI OBTENUS A PARTIR DES LISTES B CONTRE CELLES OBTENUES RESPECTIVEMENT DES LISTES B+D ET B+SA.....	20
TABLEAU 7 : CLASSEMENT SELON LES SCORES GFI DES DIFFERENTS TAUX D'ERREUR COMMIS PAR L'OMISSION DE SA DANS L'ESTIMATION DE LA FAUNE TOTALE POUR CHACUNE DES 11 RIVIERES RETENUES. LES VALEURS SUPERIEURES A 50% SONT EN ROUGE ET CELLES COMPRISES ENTRE 10% ET 50% SONT EN JAUNE	21
TABLEAU 8 : CLASSEMENT SELON LES SCORES GFI DES DIFFERENTS TAUX D'ERREUR COMMIS PAR L'OMISSION DE D DANS L'ESTIMATION DE LA FAUNE TOTALE POUR CHACUNE DES 11 RIVIERES RETENUES. LES VALEURS SUPERIEURES A 50% SONT EN ROUGE ET CELLES COMPRISES ENTRE 10% ET 50% SONT EN JAUNE	22

1 Introduction

L'évaluation écologique des grands cours d'eau fait actuellement partie des grands chantiers de la mise en œuvre de la Directive Cadre Européenne sur l'eau (DCE). Dans cette perspective, et du fait de leur grand potentiel bio-indicateur, les macro-invertébrés benthiques font partie des éléments de qualité biologique à suivre. Mais la nature de ce type de masse d'eau induit des difficultés au niveau du protocole d'échantillonnage. Le présent rapport vise à comparer les faunes en macro-invertébrés obtenues selon les trois méthodes d'échantillonnage pour l'instant mises en œuvre pour le calcul de l'Indice Biologique Global Adapté aux grands cours d'eau et aux rivières profondes (I.B.G.A., Agence de L'Eau Rhône Méditerranée Corse, Février 1997). Il est issu de l'IBGN (Indice Biologique Global Normalisé) car ce dernier est peu applicable sensu stricto aux grands cours d'eau. En effet, seule la fraction du peuplement benthique accessible à pied est concernée par ce protocole (les rives). Ceci provoque donc une sous-estimation de la qualité écologique par les calculs de variété taxonomique et du Groupe Faunistique Indicateur (GFI), car les taxons les plus polluo-sensibles se trouvent le plus souvent dans le chenal.

Jusqu'à la fin de l'année 2009, le protocole IBGA était celui qui était préconisé afin de caractériser l'état global de ce type de milieu à partir des communautés d'invertébrés benthiques. Mais à l'instar de l'IBGN, il ne correspondait pas non plus aux exigences de la DCE du fait qu'il ne prenait pas non plus en compte les peuplements de référence, les abondances, et ni les ratios entre taxons sensibles et tolérants. De plus, cette technique n'est pas comparable aux méthodes utilisées dans les autres pays européens. Il y a donc un besoin de compatibilité d'une part à la DCE et d'autre part aux méthodes appliquées pour les petits cours d'eau. Ceci afin d'assurer une continuité écologique ainsi qu'un maximum d'homogénéité dans les protocoles d'échantillonnage le long des gradients amont-aval.

Maintenant que de nombreux résultats sont disponibles, il est possible de comparer l'apport de chacune des méthodes employées pour le calcul de l'IBGA dans l'appréhension de la faune totale en macro-invertébrés. Il s'agit de proposer un outil d'aide et de réflexion à la définition du protocole intégrant les prérequis précédemment cités. Pour ce faire, cette étude est axée sur la variabilité taxonomique existant entre les méthodes de prélèvement des relevés IBGA.

Les analyses porteront directement sur le contenu des listes taxonomiques afin d'affiner les résultats jusqu'au niveau des familles.

2 Données disponibles

2.1 Bassins et rivières concernées

Les données ont été obtenues de l'Agence de l'Eau Rhône Méditerranée Corse et comprennent des résultats concernant les bassins de Loire-Bretagne (LB), de la Meuse (M), du Rhin (R), de Rhône-Méditerranée-Corse (RMC) et de Seine-Normandie (SN). Si on considère qu'un échantillon est défini par le regroupement des échantillons unitaires effectués dans un même lieu (site) et à une même date, alors on dispose de respectivement (Tableau 1) :

BASSIN	LB	M	R	RMC	SN
NOMBRE D'ECHANTILLONS	21	28	6	400	80

Tableau 1 : Table de contingence des échantillons disponibles par bassin hydrographique

Le Tableau 2 donne la même contingence des échantillons disponibles mais par cours d'eau.

Rivière	Nombre d'échantillons disponibles	Rivière	Nombre d'échantillons disponibles	Rivière	Nombre d'échantillons disponibles	Rivière	Nombre d'échantillons disponibles
LB	21	Hérault	4	Rhin (gd canal)	2	Tet	4
Ain	16	Isère	15	Rhône	178	Vesle	4
Aisne	11	Lez	2	Sane Vive	2	Vieux-Rhin	4
Aube	9	Marne	12	Saône	129	Voire	4
Aude	5	Meuse	28	Saulx	2	Yonne	8
Doubs	14	Oise	8	Seille	22		
Gard	3	Orb	6	Seine	22		

Tableau 2 : Table de contingence des échantillons disponibles par cours d'eau

Comme les rivières du bassin LB ne nous sont pas connus, les analyses se limiteront à l'échelle du bassin.

Les analyses qui se feront par rivière se justifient aisément par la grande différence entre le nombre de données disponibles par rivière, mais aussi la variabilité naturelle (climat, écorégion...) et humaine (équipes d'opérateurs différentes par bassin). Les cours d'eau avec au moins 10 échantillons ont été retenus ici et sont indiquées sur la Figure 1.

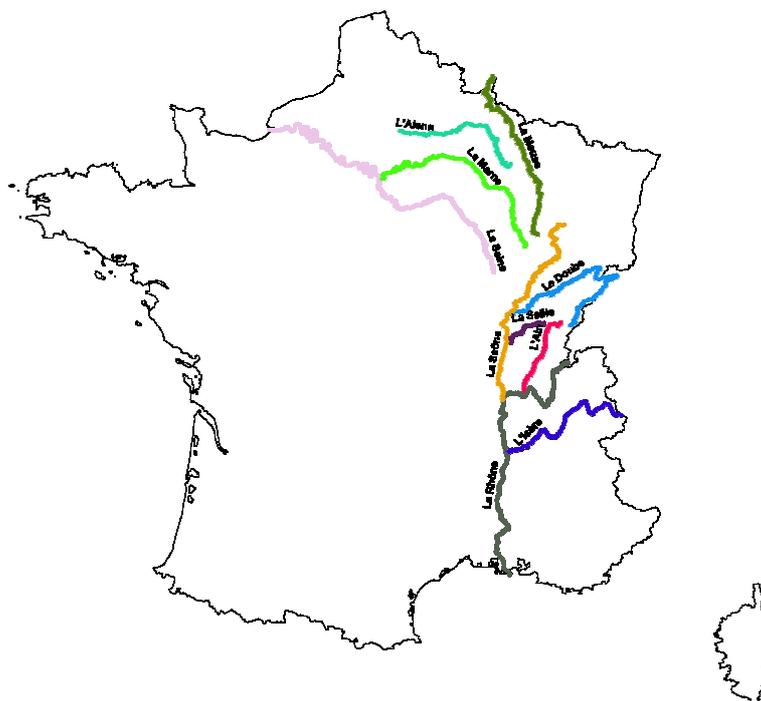


Figure 1 : Carte des cours d'eau retenus pour l'analyse du protocole d'échantillonnage de l'IBGA

2.2 Méthode d'échantillonnage

Les résultats sont issus de la mise en œuvre du protocole expérimental pour l'Indice Biologique Global Adapté aux grands cours d'eau et aux rivières profondes (I.B.G.A., Agence de L'Eau Rhône Méditerranée Corse, Février 1997). Un échantillon correspond à 13 prélèvements répartis en trois méthodes comme suit :

- 8 en rives ou Berges : il s'agit de la transposition de la méthode IBGN (ancienne version) sur les berges des cours d'eau dans leur portion praticable. Les données obtenues sont donc issues de huit échantillons unitaires.
- 3 en zone profonde par Dragage du chenal : Opérés depuis une embarcation, les dragages du chenal sont répartis sur un transect perpendiculaire aux rives. Les données obtenues sont issues des trois échantillons unitaires les plus représentatifs du chenal en termes de diversité d'habitats.
- 2 en zone profonde sur Substrat artificiel : Quatre à six substrats artificiels sont répartis sur les deux rives du cours d'eau dans des habitats présentant des caractéristiques différentes. Les données obtenues sont issues des deux échantillons unitaires les plus riches parmi tous ceux effectués pendant l'échantillonnage.

Les méso-habitats concernés par ces trois méthodes ne sont donc pas forcément les mêmes (la méthode par dragage étant par exemple la seule à être opérée dans le chenal). De plus la quantité d'information retenue par chacune des méthodes est attendue biaisée par la différence du nombre de prélèvements effectués respectivement. Ce ne sont ni la variabilité naturelle des résultats ni la variabilité intra méthode mais la variabilité inter-méthode des listes faunistiques qui sont analysées. Ainsi chaque méthode est comparée de manière équivalente dans un premier temps, puis selon une orientation rendant explicite la méthode Berge. Par commodité la majuscule B ou le terme 'Berge' font référence à la méthode sur Berges, la majuscule D ou le terme 'Drague' à la méthode par Dragage et les majuscules SA ou le terme 'Substrat Artificiel' à la méthode sur Substrat Artificiel.

2.3 Données taxonomiques

Chaque échantillon est associé à trois listes faunistiques, une par méthode. Chacune de ces listes représente la somme des abondances par taxon identifié dans chacun des échantillons unitaires d'une même méthode. Le niveau taxonomique d'identification retenu pour cette étude est la famille, ou à défaut un groupe plus élevé pour certains taxons (*i.e.* Oligochètes, Spongiaires, Bryozoaires, Cnidaires, Némertes et Nématodes). Chaque liste est donc caractérisée par un seul lieu, une seule date, et une seule méthode.

La présente analyse est effectuée sur les 535 échantillons disponibles. Comme à chaque échantillon correspond 3 méthodes de prélèvement, il y a donc 535 x 3 soit 1605 listes.

2.4 Approche comparative

Tous les cas de figure concernant un taxon donné à un site donné peuvent par exemple se résumer à partir du schéma de la synthèse additive des trois couleurs primaires (Figure 2). Ce principe est repris ici pour visualiser les associations de listes effectuées.

Le croisement des occurrences taxonomiques par méthodes correspond à celle de la juxtaposition des couleurs correspondantes. Les 8 possibilités ou événements incompatibles sont alors :

- Le taxon n'est pas trouvé dans le site alors qu'il s'y trouve (partie en gris).
- Le taxon est trouvé dans une seule et unique méthode (les trois couleurs primaires : la

partie en rouge pour la méthode sur Berges, en vert pour la méthode par Dragage et en bleu pour la méthode sur Substrat Artificiel).

- Le taxon est trouvé grâce à deux méthodes et jamais grâce à la 3^{ème} (les trois couleurs secondaires : la partie en jaune pour B + D, en cyan pour D + SA et en magenta pour B + SA).
- Le taxon est trouvé grâce aux trois méthodes (partie en blanc).

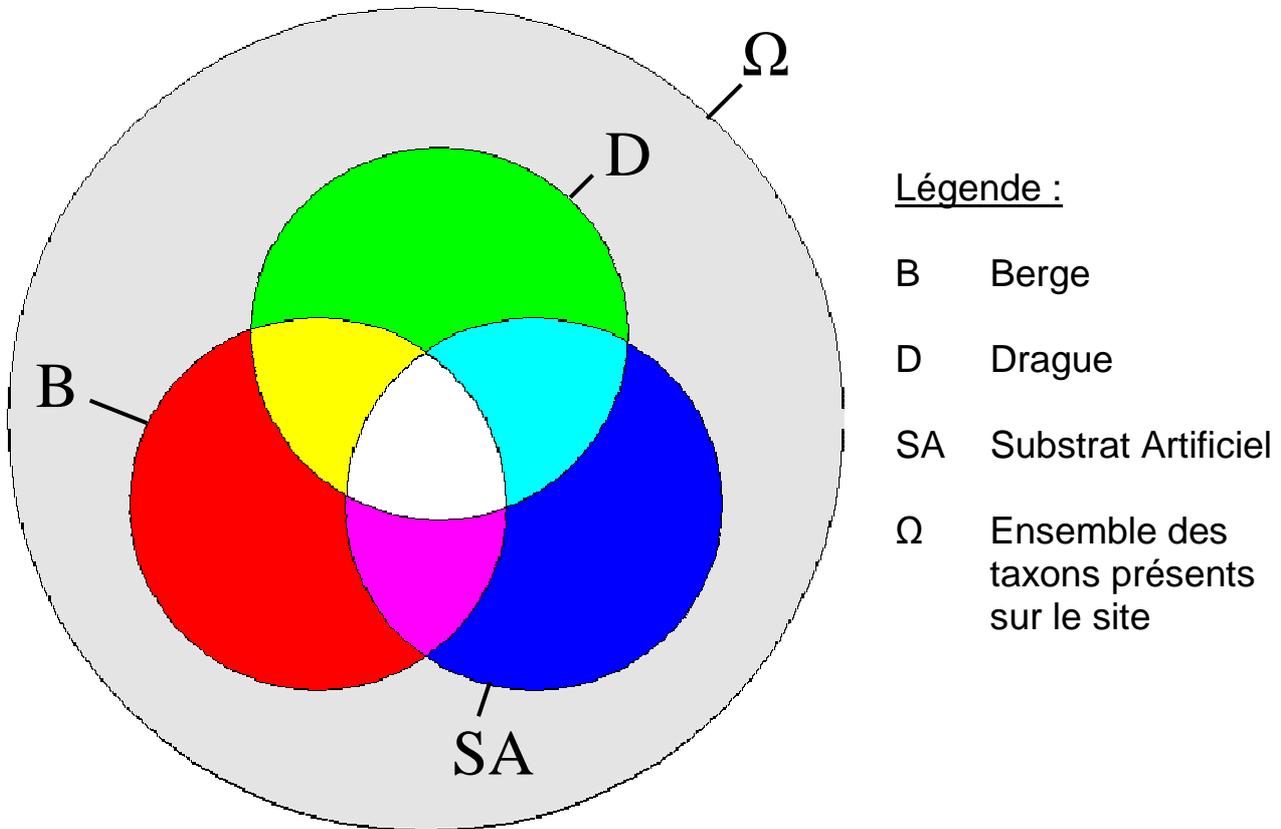


Figure 2 : Schéma regroupant les différentes possibilités d'occurrence des taxons par site parmi les différentes méthodes d'échantillonnage. La logique est la même que celle du cercle des couleurs primaires ici repris. Les couleurs primaires correspondent aux cas où un taxon est exclusivement trouvé dans une seule liste d'un site. Les couleurs secondaires correspondent aux cas où un taxon donné est exclusifs à deux méthodes d'échantillonnage. Enfin, le blanc et le noir (en gris sur le schéma par commodité graphique) correspondent respectivement à la situation où le taxon est vu dans les trois listes ou au contraire dans aucune

3 Analyses comparatives

3.1 Apports quantitatifs en taxons (richesse) selon la technique d'échantillonnage

Les richesses en familles de taxons ont été calculées pour chaque échantillon et selon chaque méthode (B, D ou SA) mais aussi selon chaque combinaison de méthode d'échantillonnage (B+D, B+SA, D+SA, et les 3).

La Figure 3 permet de visualiser les distributions des résultats regroupés par rivières. Seules les dix rivières ayant au moins 10 échantillons ont été analysées, soient respectivement : l'Ain, l'Aisne, le Doubs, l'Isère, la Marne, la Meuse, la Saône, la Seille et la Seine (Figure 1). Des tests de comparaison de moyennes appariées ont été effectués (Tableaux 4, 5 et 6). Il s'agit de tests de *t* pour données appariées avec une hypothèse alternative symétrique dans le 1^{er} tableau (pas d'*a priori* de

différence) et avec une hypothèse alternative ciblant une augmentation pour les 2 suivants car on s'attend seulement à une augmentation de richesse (sauf pour le 3^e test du Tableau 5 qui est une comparaison sans incrément de liste, pas d'*a priori* non plus). Les valeurs du risque de première espèce (p-value) sont indiquées entre parenthèse. Elles sont remplacées par des astérisques suivant leur degré de significativité, c'est à dire si elles s'avèrent inférieures aux valeurs seuils (α) de 0.05 ou 0.01.

On constate que les résultats sont assez homogènes par rivière et que la méthode sur berge est celle qui apporte le plus de richesse (moyenne des différences toujours positive et maximale). La différence entre les deux autres méthodes est moins marquée mais la méthode Substrat Artificiel semble meilleure que Dragage. Par contre si on associe la faune issue de la méthode Berge avec soit celle de la méthode Dragage ou Substrat Artificiel, les richesses combinées deviennent équivalentes ce qui laisse penser que, d'une part, la proportion de richesse de Berge est bien supérieure à celle de Dragage ou Substrat Artificiel, et que d'autre part l'apport de Dragage sur Berge est moins redondant que celui des Substrat Artificiel sur Berge.

Rivières	B Vs D Différence moyenne	B Vs SA Différence moyenne	D Vs SA Différence moyenne
Ain	13.4 (**)	7.9 (**)	-5.5 (**)
Aisne	10.3 (**)	6.8 (**)	-3.5 (0.17)
Doubs	4.4 (0.053)	3.3 (*)	-1.1 (0.518)
Isère	5.2 (**)	3.9 (*)	-1.3 (0.449)
Marne	17.1 (**)	14.4 (**)	-2.7 (0.095)
Meuse	12.1 (**)	11.5 (**)	-0.6 (0.688)
Rhône	10.3 (**)	7.1 (**)	-3.2 (**)
Saône	10.7 (**)	9.4 (**)	-1.3 (*)
Seille	12.3 (**)	9.6 (**)	-2.6 (*)
Seine	13.2 (**)	8.1 (**)	-5.1 (*)

Tableau 3 : Tests de comparaison des moyennes appariées selon la méthode d'échantillonnage entre distributions de richesses par rivière. Les p-values sont indiquées entre parenthèses. En jaune quand la p-value est significative (< 0.05), noté *. En rouge quand la p-value est très significative (< 0.01), noté **

Rivières	B Vs (B + D) Différence moyenne	B Vs (B + SA) Différence moyenne	(B + D) Vs (B + SA) Différence moyenne
Ain	-3.6 (**)	-3.3 (**)	0.3 (0.575)
Aisne	-5.5 (**)	-4.8 (**)	0.7 (0.384)
Doubs	-4.8 (**)	-5.0 (**)	-0.2 (0.812)
Isère	-5.7 (**)	-6.1 (**)	-0.3 (0.684)
Marne	-4.8 (**)	-3.7 (**)	1.1 (0.178)
Meuse	-4.0 (**)	-4.2 (**)	-0.2 (0.736)
Rhône	-4.3 (**)	-4.5 (**)	-0.2 (0.459)
Saône	-4.1 (**)	-4.0 (**)	0.1 (0.769)
Seille	-3.3 (**)	-4.5 (**)	-1.2 (*)
Seine	-4.0 (**)	-4.8 (**)	-0.8 (0.305)

Tableau 4 : Tests de comparaison des moyennes appariées selon les combinaisons de méthodes d'échantillonnage entre distributions de richesses par rivière. Les p-values sont indiquées entre parenthèses. En jaune quand la p-value est significative (< 0.05), noté *. En rouge quand la p-value est très significative (< 0.01), noté **

Rivières	(B + D) Vs Tot. Différence moyenne	(B + SA) Vs Tot. Différence moyenne
Ain	-2.6 (**)	-2.9 (**)
Aisne	-3.2 (**)	-3.9 (**)
Doubs	-3.4 (**)	-3.2 (**)
Isère	-4.1 (**)	-3.7 (**)
Marne	-1.6 (**)	-2.7 (**)
Meuse	-2.9 (**)	-2.7 (**)
Rhône	-3.0 (**)	-2.9 (**)
Saône	-2.8 (**)	-2.9 (**)
Seille	-3.3 (**)	-2.1 (**)
Seine	-3.3 (**)	-2.5 (**)

Tableau 5 : Tests de comparaison des moyennes appariées selon les combinaisons de méthodes d'échantillonnage entre distributions de richesses par rivière. Les p-values sont indiquées entre parenthèses. En jaune quand la p-value est significative (< 0.05), noté *. En rouge quand la p-value est très significative (< 0.01), noté **

L'apport de chacune des 3 méthodes d'échantillonnage augmente significativement la richesse de l'échantillon. La méthode Berge est de loin la plus riche et, de part la nature de son protocole, la plus éclectique quant à la diversité des habitats prospectés. Comme il apparaît peu envisageable de se passer de cette méthode, les analyses suivantes se focaliseront donc sur les bénéfices apportés par les deux autres méthodes en termes de représentativité sur la faune total de l'échantillon.

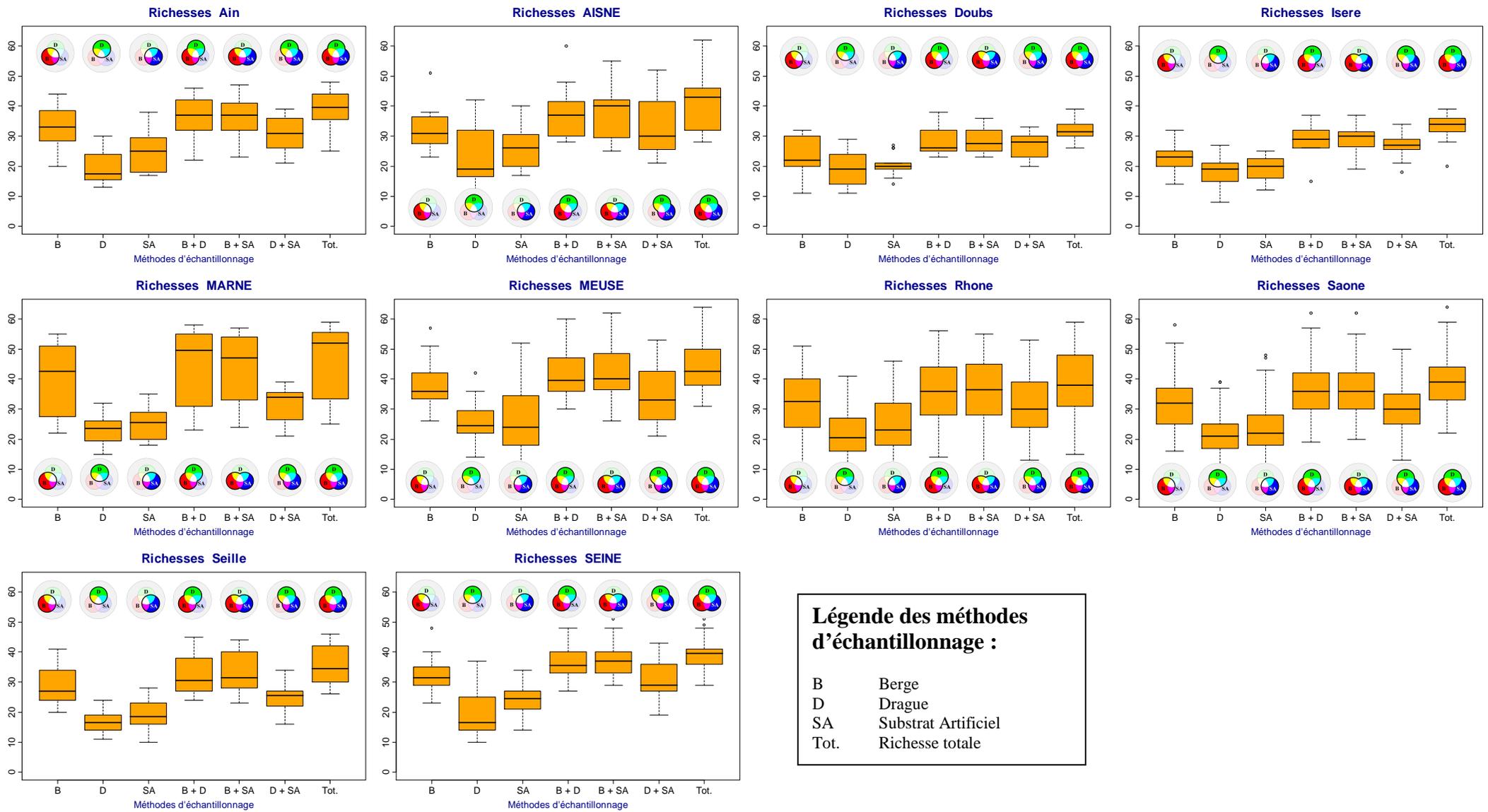


Figure 3 : Distributions des richesses en famille de macro-invertébrés par rivière et selon les méthodes d'échantillonnage prises en compte

3.2 Comparaison des distributions taxonomiques

Les données ont été préalablement triées afin d'éliminer les taxons dont l'occurrence par liste, et non par échantillon, est inférieure à 3 individus. En effet comme cette partie se concentre sur les différences de distribution en taxons entre méthodes d'échantillonnage, il est important de limiter l'effet des taxons rares ou d'occurrence accidentelle.

Dans le but de mieux appréhender les interrelations entre les listes taxonomiques de chaque méthode, des Analyses Factorielles des Correspondances (AFC) effectuées par cours d'eau permettent de les situer les unes par rapport aux autres dans l'espace factoriel. Les scores (ou projection des points sur les axes factoriels) des échantillons sont alors séparés en fonction de leur protocole. Aucun effet temporel n'a été identifié car les ellipses de dispersion regroupant les scores des listes par méthode ne sont pas ou peu influencés par les dates de prélèvements (données non montrées).

Les analyses portent sur les abondances transformées avec un logarithme naturel plus une unité afin d'éviter toute valeur négative. Les résultats pour les dix rivières ayant suffisamment de données ont été classés comme suit (les valeurs propres des 2 premiers axes sont données à la suite) :

TYPE 1 : Type « Rhône » ou pas de distinction

Pas de séparation nette par méthode (Figure 4).

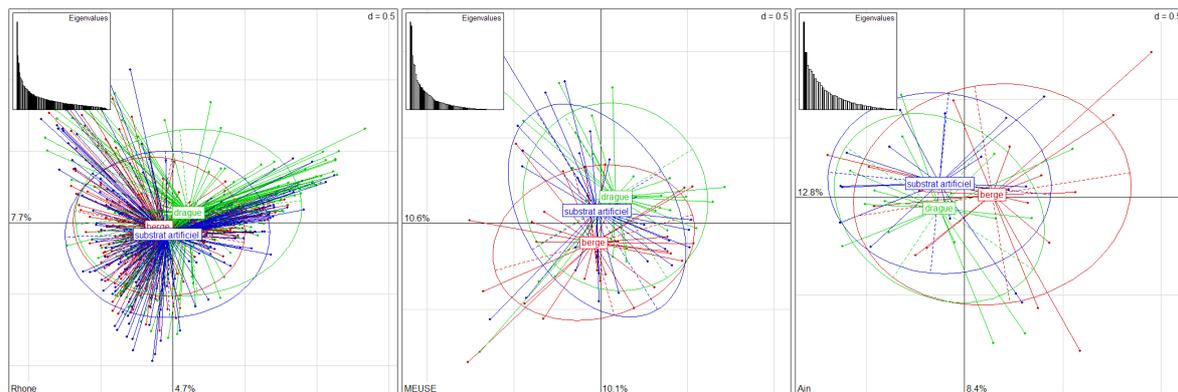


Figure 4 : Scores des listes des AFC effectuées pour les rivières du Type 1. Les scores sont regroupés par méthode d'échantillonnage. Les valeurs propres des AFC sont indiquées en proportion

Il semble pour ce type de rivière que les listes faunistiques par méthodes n'expliquent pas la majeure partie de la variabilité totale des données. Les distributions par méthode semblent donc avoir une certaine homogénéité.

TYPE 2 : Type « Saône » ou distinction d'une des méthodes

Une seule méthode se distingue des 2 autres (distinction de Dragage OU Substrat Artificiel)

- Dragage (Figure 5)

Les scores des listes issues de la méthode Dragage se distinguent assez nettement des deux autres. Par contre les ellipses de ces dernières restent imbriquées l'une dans l'autre.

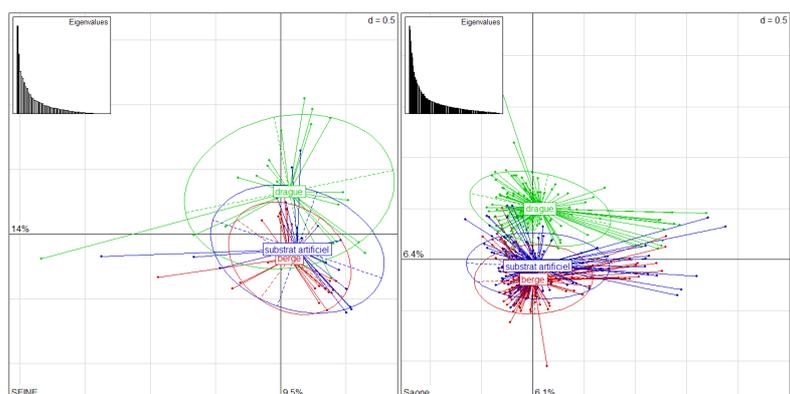


Figure 5 : Scores des listes des AFC effectuées pour les rivières du Type 2 (Drague). Les scores sont regroupés par méthode d'échantillonnage. Les valeurs propres des AFC sont indiquées en proportion de l'inertie totale

- Substrat Artificiel (Figure 6)

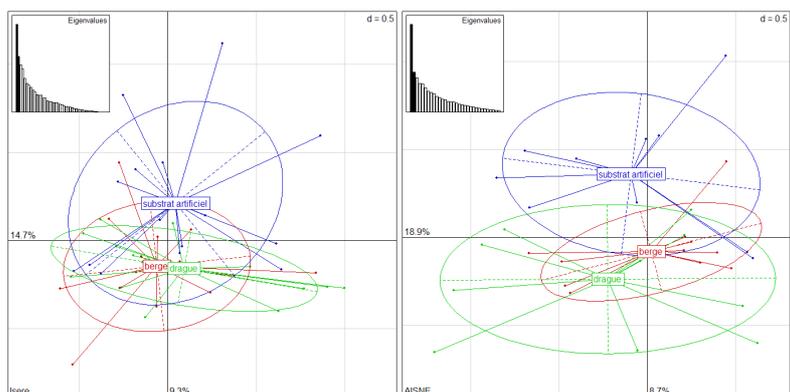


Figure 6 : Scores des listes des AFC effectuées pour les rivières du Type 2 (Substrat Artificiel). Les scores sont regroupés par méthode d'échantillonnage. Les valeurs propres des AFC sont indiquées en proportion de l'inertie totale

Ces 2 rivières possèdent moins de relevés, ce qui influe sur les résultats. En effet les regroupements de scores sont moins marqués. Néanmoins, les scores des listes issues de la méthode Substrat Artificiel se distinguent assez bien des deux autres. Par contre les centres de gravité des ellipses de ces dernières restent proches l'un de l'autre.

TYPE 3 : Type « Seille » ou Alignement des méthodes

Ce type se caractérise par un alignement des scores selon les 3 méthodes le long du 1^{er} ou 2^e axe propre, dans l'ordre absolu : B, SA, D (Figure 7).

L'ordre absolu rencontré ici tend à renforcer l'idée selon laquelle SA est plus proche de B que ne l'est D. L'effet méthode se manifeste sur une seule dimension contrairement aux autres types où les méthodes sont plutôt dispersées sur un plan que sur un axe. On peut alors supposer que dans ce type, les 3 méthodes sont plus complémentaires les unes des autres que pour les autres types.

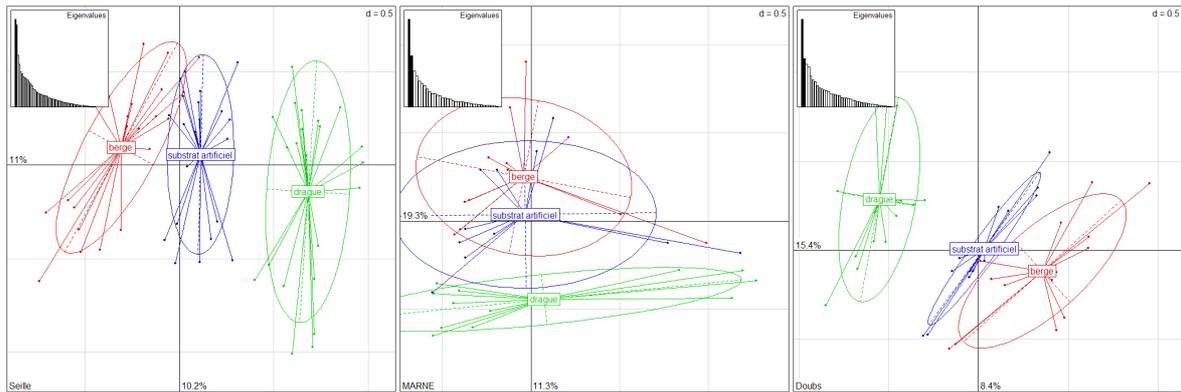


Figure 7 : Scores des listes des AFC effectuées pour les rivières du Type 3. Les scores sont regroupés par méthode d'échantillonnage. Les valeurs propres des AFC sont indiquées en proportion de l'inertie totale

Pour conclure cette partie, les rivières ont été comparées entre elles d'après les scores des listes (donc par échantillon) et non pas d'après les scores des taxons. De plus, seul le 1^{er} plan propre a été inspecté car les valeurs propres des axes suivants ne le permettent souvent pas. On retient donc qu'il existe quatre cas de figure même si ce n'est sans doute pas exhaustif car tous les grands cours d'eau de France ne sont pas utilisés ici (Figure 1). Si on compare les méthodes Drague et Substrat Artificiel à la méthode Berge, on remarque que la méthode Substrat Artificiel est à nouveau celle qui est la plus proche de la méthode Berge (sauf concernant l'Aisne et l'Isère). Néanmoins la typologie des ordinations laisse à penser qu'il ne faut pas généraliser les différences entre les 3 méthodes d'échantillonnage mais qu'il faut aussi prendre en compte les spécificités de chaque cours d'eau.

3.3 Étude des occurrences en taxons par techniques d'échantillonnage

Les données de cette partie ont été triées de la même manière que pour la partie 3.2. Les données LB correspondent au bassin de Loire Bretagne sans qu'il ait été possible d'en distinguer les cours d'eau.

3.3.1 Apport sur la liste B de la prise en compte d'une seconde technique d'échantillonnage

Les tableaux du fichier tableur « [Proportions de sites par rivière dont le taxon concerné est ajouté à la liste B grâce à la liste SA ou D.xls](#) » associe le nombre moyen d'occurrence où le taxon x est apporté à la rivière y si on ajoute soit la liste issue de la méthode Drague, soit celle issue de la méthode Substrat Artificiel à celle de la méthode Berge. Il faut noter que, pour un même échantillon, les taxons déjà présents dans la liste Berge ne sont pas pris en compte dans les résultats car c'est la contribution à la richesse faunistique totale des deux autres méthodes à cette dernière qui nous intéresse. Ceci explique pourquoi certaines valeurs sont nulles.

En comparant les valeurs moyennes de proportions de sites bénéficiant d'un apport de taxon avec la méthode Substrat Artificiel avec celles bénéficiant d'un apport de taxon de Drague (Figure 7), on constate une tendance similaire même si la méthode Substrat Artificiel semble être plus efficace quand on s'intéresse aux deux premiers taxons. Mais en ne tenant pas compte de ces deux taxons, les valeurs sont alors comparables avec celles de la méthode Drague.

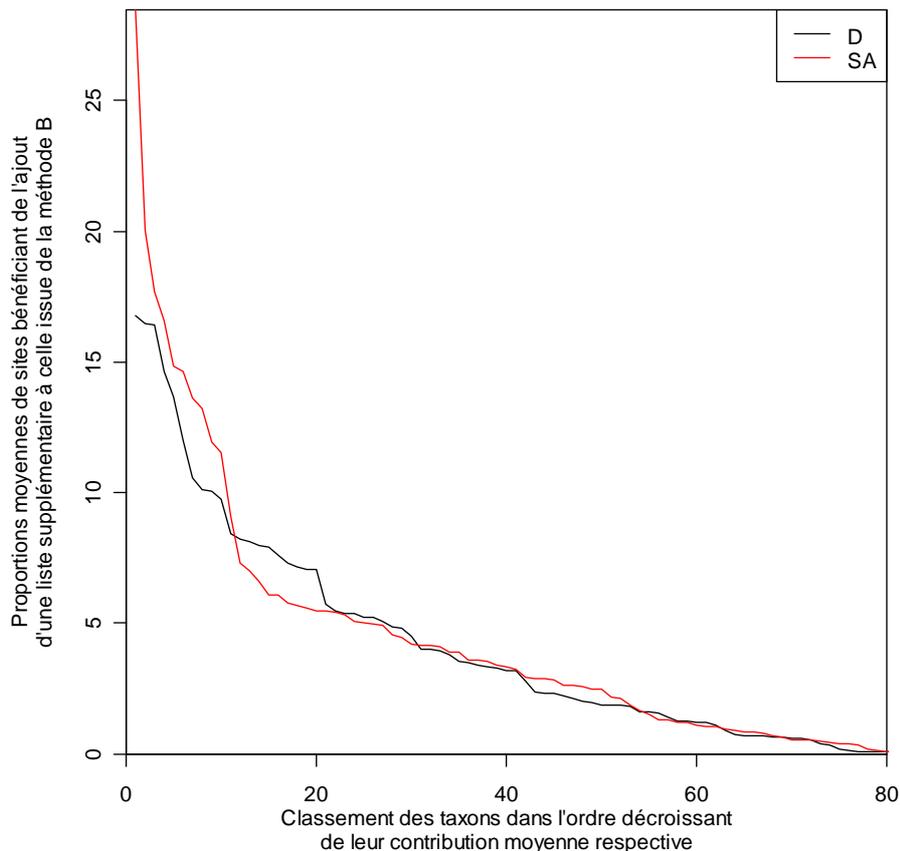


Figure 8 : Évolution des valeurs moyennes par taxon des proportions de sites bénéficiant de l'ajout d'une liste taxonomique issue soit de la méthode D (noir), soit de la méthode SA (rouge) par rapport à la liste issue de la méthode B. les valeurs sont classées dans l'ordre décroissant respectif à chaque méthode

3.3.2 Apport de la prise en compte d'une troisième technique d'échantillonnage

Afin de compléter cette approche sur les apports en taxons, les taux d'erreur auquel on s'expose si on ne pratique pas une des 3 méthodes d'échantillonnage (et plus particulièrement les méthodes Substrat Artificiel et Drague) ont été calculés.

Le principe est d'estimer le pourcentage de chance de rater un taxon donné si on n'utilise pas une des trois méthodes de prélèvement. Cela donne un « taux d'erreur » que l'on peut assimiler au taux de taxons exclusifs à une seule méthode.

La base de la comparaison se fera donc à partir des données en confrontant les résultats obtenus à partir de deux méthodes contre ceux obtenus à partir des trois méthodes ensemble. Ce dernier cas de Figure est donc supposé être ici la référence dans l'estimation de la population réelle du site évalué. L'autre hypothèse de départ est l'applicabilité de la théorie moderne des probabilités selon l'axiomatique de Kolmogorov (rappels en Annexe A). On a donc pour une rivière et un taxon donné, un ensemble d'événement supposés représentatifs et suffisants pour correspondre à la présence ou non dudit taxon selon telle ou telle méthode d'échantillonnage.

Le calcul du taux d'erreur passe par l'estimation de « l'erreur moyenne » commise si on ne prend pas en compte une des 3 méthodes d'échantillonnage. Cela revient à calculer la proportion de sites où un taxon donné est absent des 2 méthodes prises en compte dans l'évaluation de la rivière

concernée alors qu'il est trouvé d'après la 3^{ème} méthode. La traduction en langage mathématique est donc la Formule (1) reprise du théorème des probabilités composées :

$$p\left(\left(A_i \cap \sum \bar{A}_{j \neq i}\right) / \left(\cup \sum A_j\right)\right) = \frac{p\left(A_i \cap \sum \bar{A}_{j \neq i}\right)}{p\left(\left(\cup \sum A_j\right)\right)} \quad (1)$$

$p\left(A_i \cap \sum \bar{A}_{j \neq i}\right)$ est l'événement « le taxon x est trouvé dans la liste i mais pas dans les autres » et $p\left(\cup \sum A_j\right)$ est l'événement « le taxon x est présent dans l'échantillon ».

Par exemple, pour le cas où on ne prendrait pas en compte Substrat Artificiel, (1) devient :

$$p\left(\left(SA \cap \bar{B} \cap \bar{D}\right) / \left(SA \cup B \cup D\right)\right) = \frac{p\left(SA \cap \bar{B} \cap \bar{D}\right)}{p\left(SA \cup B \cup D\right)} \quad (2)$$

En appliquant (1), le numérateur de (2) devrait être $p\left(\left(SA \cap \bar{B} \cap \bar{D}\right) \cap \left(SA \cup B \cup D\right)\right)$ mais cette relation peut se simplifier en $p\left(SA \cap \bar{B} \cap \bar{D}\right)$. La démonstration est donnée en Annexe B.

D'après la Figure 2, l'exemple de l'équation (2) revient conceptuellement à faire le ratio de la surface bleue sur celle occupée par les trois cercles entrecroisés. Donc, on divise les proportions exclusives à une méthode par la proportion globale de trouver un taxon donné.

Cet exercice est réalisé pour chacune des onze rivières retenues pour l'étude. De plus une estimation par la technique de bootstrap de l'erreur standard est aussi calculée. Cette méthode peut être résumée en 6 points successifs (annexe C).

On obtient alors 3 tables reprenant chacune les différents taux d'erreur par technique et compilées dans le fichier tableur « [taux d'erreur sur 11 rivières et par taxon pour l'omission d'une des 3 techniques d'échantillonnage.xls](#) ». Les valeurs sont rangées par taxons en ligne et par rivière en colonne puis sont classées par colonne (et donc par taxon) dans l'ordre décroissant des valeurs moyennes entre rivières par taxon.

Par exemple, le résultat de la 2^{ème} colonne de la table Substrat Artificiel représente les chances que l'on a de rater les Cambaridés dans l'Ain sachant qu'il n'a pas été trouvé dans les listes Berge ni Drague. En effet, le problème de trouver ce taxon grâce à Substrat Artificiel alors qu'il est déjà trouvé grâce à Berge et/ou Drague ne se pose pas.

3.3.3 Interprétations taxonomiques

Pour une meilleure vision synthétique des résultats, la Figure 9 représente les taux d'erreur par catégorie de taxon et par rivière en histogrammes cumulés. Seuls les taxons ayant une probabilité estimée supérieure ou égale à 50% ont été pris en compte et ils ont été regroupés en six classes par simplification à savoir : les Éphémères-Plécoptères-Trichoptères (EPT), les Crustacés, les Mollusques, les Diptères, les Insectes (hormis EPT & Diptères) et le reste.

La méthode Berge apporte 60% des richesses totales moyennes (sauf pour l'Isère, 15%) et la distribution selon les grands groupes taxonomiques indiqués ci-dessus est relativement homogène. Si les apports en richesse des techniques Substrat Artificiel ou Drague sur Berge sont quantitativement comparables (partie 3.1.), on constate que la nature des taxons apportés diffère.

La méthode Substrat Artificiel représente environ 12% des richesses totales moyennes et favorise principalement l'ajout de Cambaridés et des Viviparidés. Elle permet aussi de trouver (par ordre décroissant) des Polycentropodidés, des Bithyniidés, des Asellidés, des Gomphidés, des Glossiphonidés et des Platycnemididés. Ce sont sommairement des familles ayant une valeur bio-indicatrice modérée à faible et comprennent des larves prédatrices (Odonates et Achètes).

La méthode Drague représente aussi environ 12% des richesses totales moyennes et favorise l'ajout d'Hydropsychidés, de Glossiphoniidés, d'Erpobdellidés, de Dresseinidés, de Ceratopogonidés, de Corbiculidés, d'Ephemeridés, d'Heptageniidés. Cette méthode a plutôt tendance à révéler des familles plus sensibles que celles ajoutées par la technique Substrat Artificiel.

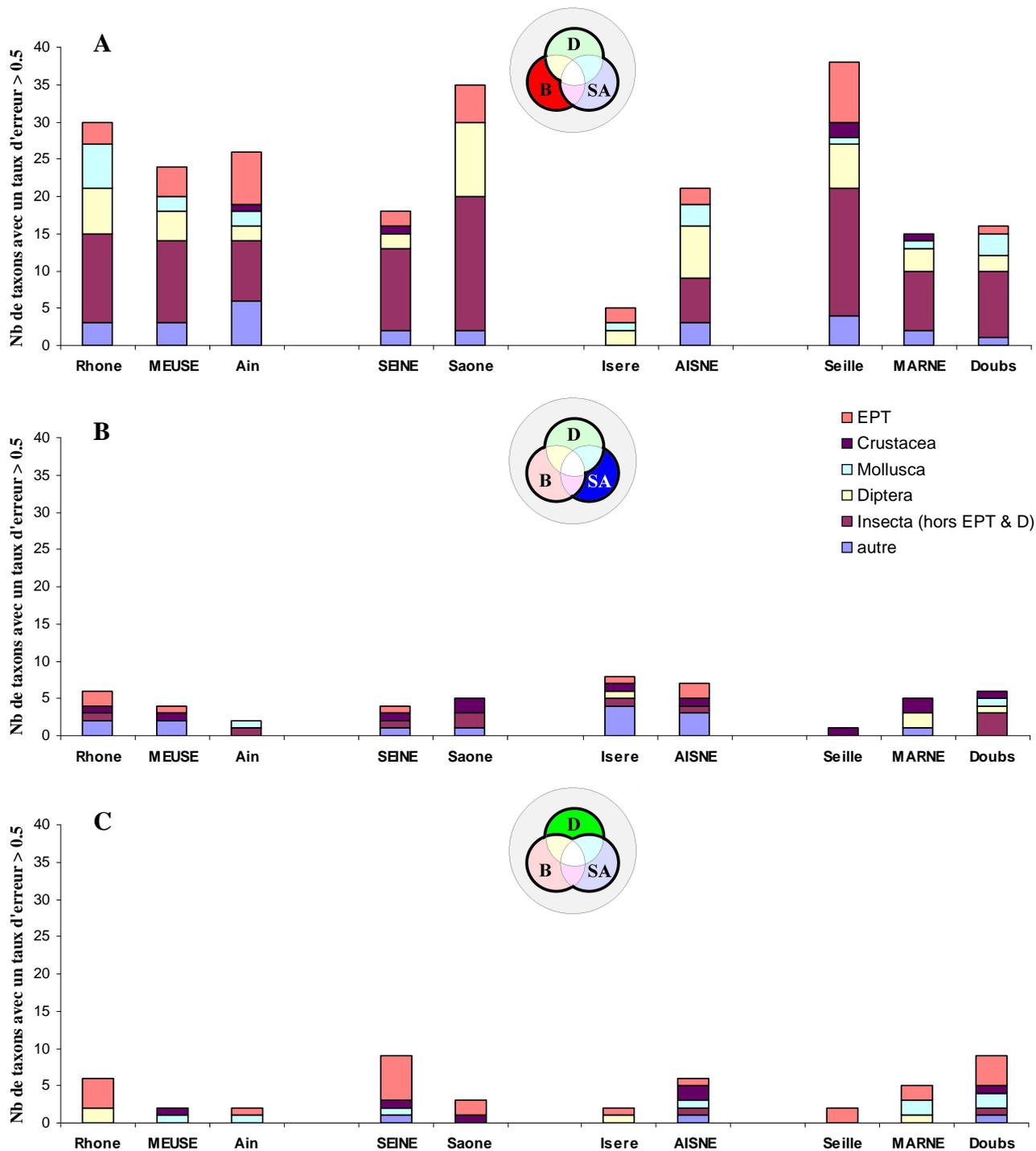


Figure 9 : Histogrammes des taux d'erreur cumulé par catégorie de taxon et par rivière selon la méthode d'échantillonnage : A- Berges, B- Substrat Artificiel et C- Drague. Les rivières sont regroupées selon les types identifiés en 3.2

3.3.4 Taxons bio-indicateurs

Pour mieux se rendre compte des différences entre les familles ajoutées à la liste issue de la méthode Berge par la méthode Substrat Artificiel ou Drague, on peut comparer les évolutions de la métrique GFI (Groupes Faunistiques Indicateurs) associées à l'indice IBGN. Un test de t pour données appariées permet de comparer les moyennes des GFI obtenus d'après la liste Berge seule avec ceux obtenus à partir des listes associant Berge avec Drague d'une part et Berge avec Substrat Artificiel d'autre part. Les résultats sont indiqués dans le Tableau 6.

Données comparées	Moyenne des écarts	p-value
GFI de B contre GFI de (B + D)	0.277	10^{-3}
GFI de B contre GFI de (B + SA)	0.01	0.887

Tableau 6 : Tests de t de comparaisons des moyennes des scores GFI obtenus à partir des listes B contre celles obtenues respectivement des listes B+D et B+SA

On voit donc que l'ajout de la liste Drague influence significativement la valeur du GFI alors que l'ajout de la liste Substrat Artificiel n'apporte pas grand chose de ce point de vue là.

Les valeurs des taux d'erreurs auquel on s'expose si on ne pratique pas Substrat Artificiel ou Drague ont été respectivement reprises pour les taxons utilisés dans le calcul de la métrique GFI (ou taxons GFI). De plus, elles ont été reclassées dans le même ordre que la grille de la métrique GFI, soit par ordre décroissant des groupes faunistiques indicateurs. Les valeurs concernant les taux d'erreur Substrat Artificiel (tableau 7) des taxons GFI sont plus dispersées et moins nombreuses que celles concernant les taux d'erreur Drague (tableau 8). Cette dernière semble influente sur les résultats en note GFI surtout pour les scores compris entre 3 et 7. Pour les scores GFI supérieurs à 7, les résultats sont plutôt propres à une ou deux rivières seulement. Par exemple, les Chloroperlidés ne sont manqués que dans le Rhône si on ne prend pas en compte les résultats de la méthode Substrat Artificiel. Les taxons de ces catégories sont souvent assez rares et il ne faut pas oublier que les données ont été triées en éliminant les faibles occurrences. Par contre, pour les scores GFI inférieurs ou égaux à 7, les résultats sont plus généraux car impliquant plusieurs cours d'eau. Par exemple, les Leuctridés sont manqués dans toutes les rivières hormis la Meuse et le Bassin Loire-Bretagne si on ne prend pas en compte les résultats de la méthode Drague.

Taux d'erreur SA	GFI	LB	Ain	Aisne	Doubs	Isère	Marne	Meuse	Rhône	Saône	Seille	Seine
Chloroperlidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00
Perlidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Perlodidae	9	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00
Taeniopterygidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Capniidae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Brachycentridae	8	0.00	0.00	50.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00
Odontoceridae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Philopotamidae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Leuctridae	7	0.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00	0.00	0.00	0.00
Glossosomatidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	25.00	0.00	0.00	0.00
Beraeidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00
Goeridae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Leptophlébiidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nemouridae	6	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00
Lepidostomatidae	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sericostomatidae	6	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ephemeridae	6	0.00	0.00	14.29	33.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Hydroptilidae	5	16.67	0.00	0.00	0.00	0.00	9.09	4.35	8.00	3.19	6.25	0.00
Heptageniidae	5	50.00	0.00	0.00	14.29	0.00	0.00	33.33	2.17	0.00	0.00	33.33
Polymitarcidae	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Potamanthidae	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.63	33.33	0.00	0.00
Leptoceridae	4	7.69	0.00	0.00	0.00	0.00	0.00	0.00	6.12	1.32	0.00	0.00
Polycentropodidae	4	50.00	8.33	9.09	8.33	0.00	40.00	19.23	15.48	22.22	12.50	38.46
Psychomyiidae	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.21	4.76	0.00	0.00
Rhyacophilidae	4	0.00	0.00	0.00	0.00	44.95	0.00	16.67	40.00	0.00	0.00	0.00
Limnephilidae	3	0.00	0.00	0.00	0.00	11.11	0.00	75.00	12.50	0.00	0.00	0.00
Ephemerellidae	3	0.00	0.00	0.00	0.00	30.00	0.00	0.00	3.23	0.00	0.00	0.00
Hydropsychidae	3	20.00	16.67	0.00	7.69	14.29	0.00	7.41	7.58	33.33	0.00	11.11
Aphelocheiridae	3	0.00	0.00	0.00	0.00	0.00	0.00	6.67	0.00	0.00	0.00	0.00
Baetidae	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.83	1.23	0.00	0.00
Caenidae	2	4.76	12.50	0.00	0.00	0.00	8.33	0.00	8.08	0.91	11.11	0.00
Elmidae	2	0.00	0.00	0.00	0.00	55.56	28.57	0.00	6.00	2.27	0.00	0.00
Gammaridae	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.26	5.26
Mollusques	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chironomidae	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Asellidae	1	5.26	0.00	10.00	23.08	66.67	9.09	0.00	16.54	5.88	10.00	0.00
Achètes	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Oligochètes	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tableau 7 : Classement selon les scores GFI des différents taux d'erreur commis par l'omission de SA dans l'estimation de la faune totale pour chacune des 11 rivières retenues. Les valeurs supérieures à 50% sont en rouge et celles comprises entre 10% et 50% sont en jaune

Taux d'erreur D	GFI	LB	Ain	Aisne	Doubs	Isère	Marne	Meuse	Rhône	Saône	Seille	Seine
Chloroperlidae	9	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00
Perlidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Perlodidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Taeniopterygidae	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00
Capniidae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00
Brachycentridae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Odontoceridae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Philopotamidae	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00
Leuctridae	7	0.00	33.33	26.79	100.00	20.00	100.00	0.00	46.15	100.00	100.00	100.00
Glossosomatidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00
Beraeidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Goeridae	7	0.00	0.00	0.00	0.00	0.00	0.00	50.00	0.00	33.33	0.00	0.00
Leptophlébiidae	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nemouridae	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lepidostomatidae	6	0.00	33.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sericostomatidae	6	0.00	14.29	0.00	0.00	0.00	0.00	0.00	16.67	0.00	0.00	100.00
Ephemeridae	6	25.00	100.00	14.29	33.33	0.00	50.00	36.36	54.55	43.36	12.50	12.50
Hydroptilidae	5	0.00	33.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.09
Heptageniidae	5	0.00	0.00	0.00	28.57	14.29	50.00	0.00	46.76	41.18	0.00	33.33
Polymitarcidae	5	0.00	0.00	50.00	66.67	0.00	100.00	0.00	100.00	88.89	100.00	100.00
Potamanthidae	5	0.00	0.00	33.33	27.27	0.00	50.00	11.11	7.89	0.00	0.00	0.00
Leptoceridae	4	7.69	0.00	0.00	0.00	0.00	25.00	0.00	42.22	5.41	0.00	62.50
Polycentropodidae	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.71	0.00	0.00	0.00
Psychomyiidae	4	0.00	20.00	0.00	0.00	57.14	25.00	18.75	11.96	2.44	33.33	0.00
Rhyacophilidae	4	0.00	0.00	0.00	0.00	18.18	0.00	0.00	8.33	0.00	0.00	0.00
Limnephilidae	3	0.00	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00	0.00
Ephemerellidae	3	0.00	12.50	0.00	0.00	10.00	0.00	0.00	6.45	11.11	0.00	0.00
Hydropsychidae	3	0.00	33.33	22.22	15.38	0.00	50.00	3.70	15.91	10.00	0.00	50.00
Aphelocheiridae	3	0.00	0.00	0.00	66.67	0.00	0.00	33.33	0.00	0.00	0.00	50.00
Baetidae	2	0.00	8.33	0.00	0.00	0.00	0.00	0.00	6.56	0.00	0.00	0.00
Caenidae	2	0.00	25.00	0.00	7.14	0.00	0.00	0.00	9.09	0.00	5.56	0.00
Elmidae	2	28.57	6.67	0.00	0.00	0.00	0.00	7.14	18.92	11.36	0.00	23.53
Gammaridae	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.14	0.00	0.00	5.00
Mollusques	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chironomidae	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Asellidae	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.25	1.96	0.00	0.00
Achètes	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Oligochètes	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tableau 8 : Classement selon les scores GFI des différents taux d'erreur commis par l'omission de D dans l'estimation de la faune totale pour chacune des 11 rivières retenues. Les valeurs supérieures à 50% sont en rouge et celles comprises entre 10% et 50% sont en jaune

4 Discussion

4.1 Cohérence des résultats

Il s'agit d'inférer des probabilités conditionnelles plus ou moins complexes à partir des listes disponibles. Or il ne faut pas oublier que seule la nature géographique des données est connue. On suppose alors que l'état écologique des points de prélèvements n'affecte pas les résultats par exemple en faisant baisser la probabilité d'occurrence d'un taxon sensible parce que la moitié des échantillons sont perturbés. Le fait de diviser les proportions calculées par la proportion globale de trouver un taxon donné en 2.3.2. diminue ce biais.

Cas des taxons rares :

Les taxons rares et à forte valeur bio-indicatrice comme les Perlodidés ou les Chloroperlidés peuvent présenter des résultats contradictoires. Mais il ne faut pas oublier que l'étude se base à partir de la partie 2.3. sur des données triées en ne retenant que les taxons ayant une occurrence par liste supérieure à trois individus. Il est aussi fort possible que le biais expérimental joue un rôle important. Il est donc souhaitable pour ce cas de figure de faire très attention aux conclusions à tirer des résultats présentés ici, notamment en ce qui concerne leur absence des listes Berge. Aussi des conclusions approfondies à propos des taxons rares nécessitent d'autres analyses.

4.2 Différences entre les méthodes d'échantillonnage

On constate que les trois techniques restent complémentaires les unes des autres. Il y a à chaque fois une forte composante commune mais aussi une composante propre à chaque technique d'échantillonnage. La nature des taxons concernés peut varier d'une rivière à l'autre. Néanmoins, cette complémentarité se conserve plus ou moins à chaque fois.

La méthode Berge reste la méthode de référence pour diverses raisons, il est donc plutôt question de discuter des deux autres comme d'un soutien à la première. Cette discussion se base sur les résultats obtenus à partir des données.

Il faut noter que la méthode sur Substrat Artificiel est un peu plus proche de la méthode Berge que ne l'est la méthode Dragage. D'un point de vue quantitatif, le nombre moyen de taxons nouveaux apportés par ces deux techniques reste équivalent, bien que légèrement favorable à la méthode Substrat artificiel. Par contre, comme il l'a déjà été dit, les deux cohortes de taxons nouveaux sont très peu semblables. Ainsi, la méthode sur Substrat Artificiel est plutôt utile pour repérer des taxons avec une moindre polluo-sensibilité ainsi que quelques Odonates, Sialidés et Polycentropodidés. Par contre, la méthode par Dragage permet de trouver des taxons avec un panel de qualités bioindicatrices plus étendu et recouvrant par exemple des Plécoptères, des Éphéméroptères, des Mollusques, des Diptères, des Hirudinées...

Il faut enfin remarquer que le nombre et les proportions de taxons 'manqués' en ne pratiquant pas la technique sur Substrat Artificiel est un peu plus faible que si on ne pratique pas la méthode par Dragage. Cela est certainement dû au fait que les substrats artificiels sont posés sur des habitats plus proches de ceux ciblés par la méthode sur Berges que de ceux ciblés par la méthode par Dragage.

5 Conclusion

Il existe bien des différences entre chacune des méthodes justifiant leur importance respective sachant que la méthode sur Berges est toujours qualitativement et quantitativement la plus significative. En général coupler la méthode sur Berges avec la méthode par Dragage est une réelle plus value car cette dernière apporte des taxons polluo-sensibles qui ne seront pas forcément prélevés avec la première méthode. Par contre, la méthode sur Substrat Artificiel permet plutôt de palier des éventuels manques à la méthode sur Berges, même si elle peut éventuellement comprendre des taxons bio-indicateurs.

Cette constatation globale est moins évidente à l'échelle d'une seule rivière où on a remarqué des nuances quant à l'apport faunistique des listes Drague et Substrat Artificiel. Il est donc important de tenir compte du contexte environnemental afin de déterminer quelles méthodes permettront d'obtenir une image la plus représentative possible d'un grand cours pour un coût en effort d'échantillonnage le plus faible possible

Une mise en place d'un protocole préliminaire pour l'élaboration d'un indice de qualité des grands cours d'eau à partir de la faune en macro-invertébrés doit donc intégrer cette notion.

Annexes

ANNEXE A. Rappels de probabilités

Terminologie

$p(A)$ signifie probabilité que l'événement A se réalise.

$p(\bar{A})$ signifie probabilité que l'événement A ne se réalise pas (indicateur logique NON).

\cup indique qu'au moins un des événements se réalise (indicateur logique OU).

\cap indique que tous les événements se réalisent simultanément (indicateur logique ET).

D'après <http://rfv.insa-lyon.fr/~jolion/STAT/poly.html> :

Axiomatique de Kolmogorov

A chaque évènement, on associe un nombre positif compris entre 0 et 1, sa probabilité. Afin d'éviter toute discussion sur cette notion, la théorie moderne des probabilités repose sur l'axiomatique suivante :

Définition 1

On appelle probabilité sur (Ω, \mathcal{P}) (où Ω est l'ensemble des évènements et \mathcal{P} une classe de parties de Ω), ou loi de probabilité, une application P de \mathcal{P} dans $[0,1]$ telle que :

- $P(\Omega) = 1$

- pour tout ensemble dénombrable d'évènements incompatibles A_1, A_2, \dots, A_n on a $P(\cup A_i) = \sum P(A_i)$.

Définition 2

On appelle espace probabilisé le triplé (Ω, \mathcal{P}, P)

Une loi de probabilité n'est donc rien d'autre qu'une mesure positive de masse totale 1. On peut donc relier la théorie des probabilités à celle de la mesure.

Propriétés élémentaires

De l'axiomatique de Kolmogorov, on peut déduire les propriétés suivantes :

Propriété 1 : $P(\emptyset) = 0$

Propriété 2 : $P(\bar{A}) = 1 - P(A)$

Propriété 3 : $P(A) \leq P(B)$ si $A \subset B$

Propriété 4 : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Propriété 5 : $P(\cup A_i) \leq \sum_i P(A_i)$ (Il n'y a stricte égalité que si les évènements A_i sont deux à deux incompatibles.)

Propriété 6 : Continuité monotone séquentielle. Soient $A_1 \supset A_2 \supset \dots \supset A_n \supset \emptyset$

Si $\lim_{n \rightarrow \infty} A_n = \emptyset$ alors $\lim_{n \rightarrow \infty} P(A_n) = 0$

Propriété 7 : Théorème des probabilités totales : Soit $\Omega = \bigcup B_i$ un système complet d'évènements (i.e. tel que $\{B_i\}$ constitue une partition de Ω). $\forall A : P(A) = \sum_i P(A \cap B_i)$

Remarque : $P(A) = 0 \Rightarrow A = \emptyset$. De même, $P(A) = 1 \Rightarrow A = \Omega$.

Théorème des probabilités composées

Soient deux évènements A et B réalisés respectivement n et m fois au cours de N épreuves. On a donc $P(A) = \frac{n}{N}$ et

$$P(B) = \frac{m}{N}. \text{ Si de plus A et B sont réalisés simultanément } k \text{ fois, on a } P(A \cap B) = \frac{k}{N}.$$

On en déduit que La probabilité de l'évènement B sachant que l'évènement A est réalisé est appelée **probabilité conditionnelle de B sachant A** et se note $P(B/A)$. Dans notre cas, on a $P(B/A) = \frac{k}{n}$.

Par définition, on a alors $P(B/A) = \frac{P(A \cap B)}{P(A)}$ et $P(A/B) = \frac{P(A \cap B)}{P(B)}$.

ANNEXE B. Démonstration de la simplification de la formule de probabilités

Soit la formule (1) reprise du théorème des probabilités composées :

$$p(A/B) = \frac{p(A \cap B)}{p(B)} \quad (\text{B.1})$$

Si A est l'événement « le taxon x est trouvé dans la liste i mais pas dans les autres » (donc $p(A_i \cap \sum \bar{A}_{j \neq i})$) et B est l'événement « le taxon x est présent dans l'échantillon » (donc $p(\cup \sum A_j)$), alors l'équation (B.1) devient :

$$p\left(\frac{A_i \cap \sum \bar{A}_{j \neq i}}{\cup \sum A_j}\right) = \frac{p\left((A_i \cap \sum \bar{A}_{j \neq i}) \cap (\cup \sum A_j)\right)}{p(\cup \sum A_j)} \quad (\text{B.2})$$

Le numérateur de (B.2) peut se simplifier. Pour l'expliquer, il faut utiliser la propriété 7 de l'annexe A :

$$p(A) = p(A \cap B) + p(A \cap \bar{B}) \quad (\text{B.3})$$

car $p(B) + p(\bar{B}) = 1$. On en déduit alors que

$$p(A_i \cap \sum \bar{A}_{j \neq i}) = p\left((A_i \cap \sum \bar{A}_{j \neq i}) \cap (\cup \sum A_j)\right) + p\left((A_i \cap \sum \bar{A}_{j \neq i}) \cap (\cup \sum \bar{A}_j)\right) \quad (\text{B.4})$$

Or le second membre de l'addition vaut 0 car il est impossible d'avoir à la fois l'événement « seulement dans la liste i » et l'événement « dans aucune liste ». On en déduit donc que :

$$p\left((A_i \cap \sum \bar{A}_{j \neq i}) \cap (\cup \sum A_j)\right) = p(A_i \cap \sum \bar{A}_{j \neq i}) \quad (\text{B.5})$$

Et donc, l'équation (B.2) devient :

$$p\left(\frac{A_i \cap \sum \bar{A}_{j \neq i}}{\cup \sum A_j}\right) = \frac{p(A_i \cap \sum \bar{A}_{j \neq i})}{p(\cup \sum A_j)} \quad (\text{B.6})$$

ANNEXE C. Principe de l'estimation par bootstraps de l'erreur standard faite sur l'estimation des taux d'erreur

- 1- On dispose d'un échantillon de taille n de valeurs par taxon et par rivière, ici la présence ou non d'un taxon donné et dans chacune des listes issues des trois méthodes d'échantillonnage. Les triplets de résultats par méthode pour un même échantillon sont conservés afin de ne pas dissocier leurs inter-relations.
- 2- On veut la statistique s ou « erreur moyenne » qui correspond à la formule (1).
- 3- On tire un nombre suffisant $B = 1000$ de sous-échantillons x^* de taille n , constitués en tirant avec remise n valeurs dans l'échantillon de départ. Ce sont les « bootstraps ».
- 4- On calcule la statistique « erreur moyenne » pour chacun des ces bootstraps. On dispose alors de B valeurs $s(x^*)$ issues chacune d'un ré-échantillonnage.
- 5- On calcule simplement la moyenne $s(\cdot)$ de ces B valeurs.
- 6- L'estimation par bootstraps de l'erreur standard est alors calculée avec la formule (C.1):

$$s\hat{e}_{boot} = \sqrt{\frac{\sum_{i=1}^B (s(x^*) - s(\cdot))^2}{B-1}} \quad (C.1)$$

On divise par $B - 1$ (et non seulement B) afin de limiter l'effet d'une sur-estimation de l'erreur standard faite avec cette méthode.