



HAL
open science

Les essais interlaboratoires en hydrobiologie : état des lieux, propositions

N. Dagens, Christian Chauvin

► **To cite this version:**

N. Dagens, Christian Chauvin. Les essais interlaboratoires en hydrobiologie : état des lieux, propositions. [Rapport de recherche] irstea. 2013, pp.44. hal-02601856

HAL Id: hal-02601856

<https://hal.inrae.fr/hal-02601856>

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.



LES ESSAIS INTERLABORATOIRES EN HYDROBIOLOGIE

ETAT DES LIEUX, PROPOSITIONS

**Action I-A-04 - Méthodes de mesures
hydrobiologiques et transferts aux opérateurs**

Nina Dagens, Christian Chauvin

Irstea

Décembre 2013

Programme scientifique et technique

Année 2012

En partenariat avec



Avec le soutien de



et de



Contexte de programmation et de réalisation

Dans le cadre des actions Aquaref, l'Irstea est impliqué dans la démarche qualité en hydrobiologie et a mené une réflexion sur les essais interlaboratoires et notamment les essais d'aptitude sur les différentes méthodes hydrobiologiques.

Ce rapport, comprenant une analyse des essais interlaboratoires existants ainsi que des propositions d'éléments techniques pour la constitution de ces essais, a été réalisé dans le cadre du programme d'activité AQUAREF pour l'année 2012.

Auteur (s) :

Dagens Nina
Irstea - Centre de Bordeaux
nina.dagens@irstea.fr

Christian Chauvin
Irstea - Centre de Bordeaux
christian.chauvin@irstea.fr

Vérification du document :

Bénédicte Lepot
Ineris
benedicte.lepot@ineris.fr

Jean-Philippe Ghestem
BRGM
jp.ghestim@brgm.fr

Laurence Miossec
Ifremer
laurence.miossec@ifremer.fr

Les correspondants

Onema : Yorick Reyjol, yorick.reyjol@onema.fr
Emilie Breugnot, emilie.breugnot@onema.fr

Etablissement - Irstea : Christian Chauvin, christian.chauvin@irstea.fr

Référence du document : Nina Dagens, Christian Chauvin - Les essais interlaboratoires en hydrobiologie - Rapport AQUAREF 2012 - 44 p.

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

Sommaire

1.	Introduction.....	7
2.	Qu'est ce qu'un Essai Interlaboratoires ?.....	8
3.	Spécificité des méthodes hydrobiologiques	9
3.1.	Le protocole d'acquisition de la donnée dans la chaîne d'évaluation.....	9
3.2.	Travailler sur des taxons et des peuplements.....	11
3.2.1.	Ecorégionalisation	11
3.2.2.	Méthodes semi-destructives	11
3.3.	Phases techniques dans les protocoles hydrobiologiques.....	12
4.	Contenu d'un essai d'aptitude en hydrobiologie.....	13
4.1.	Les référentiels applicables	13
4.1.1.	NF EN ISO/CEI 17043 d'avril 2010	14
4.1.2.	NF ISO 13528 de décembre 2005.....	14
4.1.3.	NF EN 16101 de décembre 2012	14
4.1.4.	PrEN 14184 (mars 2012, version finale en cours d'approbation).....	15
4.2.	Quelles phases méthodologiques dans un essai d'aptitude ?.....	15
4.3.	Comparer des analyses semi-destructives	16
4.4.	Problématique de régionalisation	17
4.5.	Conditions de réalisation	18
4.6.	Adapter les essais à chaque méthode.....	18
5.	Analyse des résultats.....	22
5.1.	Données et méthode	23
5.2.	Détermination de la valeur assignée : NF ISO 13528 et NF EN ISO/CEI 17043....	24
5.3.	Calcul des statistiques de performances	24
5.3.1.	NF ISO 13528 et NF EN ISO/CEI 17043.....	24
5.3.2.	Stribling et al. (2003).....	26
5.3.3.	Définition des incertitudes - Protocole Invertébrés en cours d'eau	27
5.3.4.	Autres types de calculs.....	27
5.4.	Evaluation des performances	29
5.5.	BILAN.....	29
6.	Exemples d'essais d'aptitude existants.....	31
6.1.	Essais en cours d'eau	31
6.1.1.	Invertébrés : EIL Aglae	31

6.1.2. Diatomées : essai collaboratif des DREAL	33
6.1.3. Diatomées : comparaison interlaboratoire AGLAE	34
6.2. Essais en plans d'eau	34
6.2.1. Phytoplancton	34
6.3. Essais en eaux littorales.....	35
6.3.1. Macro-algues.....	36
6.3.2. Phytoplancton	37
7. Synthèse, perspectives.....	38

Liste des figures

- *Figure 1 : Objectifs des CIL ou EIL - source LNE.*
- *Figure 2 : La chaîne d'évaluation en hydrobiologie.*
- *Figure 3 : Importance relative de chaque étape dans les différents protocoles hydrobiologiques utilisés dans les réseaux de surveillance DCE.*
- *Figure 4 : Intégration des différentes phases de l'acquisition de la donnée aux essais d'aptitude selon les méthodes hydrobiologiques.*
- *Figure 5 : Bilan des méthodes d'analyse des résultats d'un essai d'aptitude en hydrobiologie.*
- *Figure 6 : Objectifs des EIL et des essais collaboratifs.*

RESUME :

La maîtrise de la qualité des données de surveillance des masses d'eau, en application de la Directive européenne sur l'eau, s'appuie en France sur les processus d'accréditation et d'agrément des laboratoires. Dans ces démarches, les essais interlaboratoires (EIL) sont désormais préconisés ou imposés. En hydrobiologie, les EIL étant jusqu'à ce jour peu développés, il est apparu nécessaire de mieux définir ce que peuvent être ces comparaisons, notamment en ce qui concerne leur contenu et l'analyse des résultats obtenus.

Ce rapport précise la définition ainsi que les objectifs des EIL et en particulier des essais d'aptitude, d'un point de vue général. Les spécificités des méthodes hydrobiologiques sont ensuite détaillées, ce qui permet de définir les caractéristiques des méthodes par rapport à l'application des essais. L'adaptation des essais d'aptitudes en fonction des méthodes hydrobiologiques apparaît primordiale. Les particularités méthodologiques des protocoles de mesure hydrobiologique soulèvent des questions concrètes ou conceptuelles pour la réalisation de ces essais. Dans la majorité des cas, les phases d'échantillonnage et de laboratoire ne doivent pas être dissociées. Un essai ne contenant qu'une de ces phases ne permet pas une évaluation pertinente des capacités des laboratoires à appliquer une méthode dans son ensemble, ni à fournir des résultats finaux conformes et homogènes. Il peut toutefois apporter des informations comparatives sur certains aspects des compétences de l'opérateur.

Un bilan des méthodes d'analyse des résultats d'EIL est également présenté et leur application aux essais d'aptitude en hydrobiologie est discutée. Il est important que l'évaluation des laboratoires soit faite sur la production de résultats bruts correspondant généralement à une liste taxonomique (donnée qualitative) assortie des abondances associées à chaque taxon (données quantitatives).

Ce rapport contient également un inventaire succinct de quelques essais existants, dont les objectifs et caractéristiques sont rappelés.

Des propositions cadre de contenu des essais selon les méthodes hydrobiologiques sont formulées suite à l'analyse de ces divers aspects. Ce sont les essais collaboratifs, bien qu'informels, qui paraissent répondre le mieux aux besoins actuels, en apportant des éléments de réponses aux questions fondamentales de la mise en œuvre des méthodes hydrobiologiques, concernant les compétences, la comparaison des laboratoires, la formation et l'harmonisation des pratiques. Les différents exercices existants doivent néanmoins être améliorés pour répondre pleinement aux attentes des donneurs d'ordre.

Mots clés (thématique et géographique) :

Essai interlaboratoires (EIL) - Essai d'aptitude - Hydrobiologie - Essai collaboratif - Opérateur - Laboratoires - Echantillonnage - Analyse

ABSTRACTS :

With the implementation of the European Water Framework Directive, the control of water quality monitoring data is based on the accreditation and the agreement of laboratories. Inter-laboratory comparisons (ILC) are now required by these processes. To date, inter-laboratory tests are poorly developed in hydrobiology, then it seemed necessary to define more specifically what inter-laboratory comparisons must be, particularly concerning the content of these tests and the analysis of the results.

This report specifies definitions and objectives of inter-laboratory comparisons and proficiency tests generally speaking. The specificities of each hydrobiological method allow to make the link between methods features and the application of these tests. It seems essential to fit proficiency tests to each hydrobiological method. Specificities of these measurement protocols raise concrete and conceptual questions about the execution of these ILC. Generally, sampling and laboratory phases could not be separated. When an ILC takes into account only one of these methodological stages, it will not allow a relevant assessment of laboratory skills to apply a method as a whole or to provide suitable final results. However, it can bring information to compare some specific operator skills.

The different methods used to analyse inter-laboratory comparisons results are presented and the application of these methods to an ILC in hydrobiology is discussed. Laboratory evaluation should be carried out on production of raw results, which usually are taxonomic lists (qualitative data) and abundance related to taxa (quantitative data).

An overview of the existing inter-laboratory comparisons and ring-tests can also be found in this report and reminds features and objectives of these tests.

A content of inter-laboratory comparison according to hydrobiological methods is then proposed. Indeed, collaborative trials, even if they are informal, appear as the better way for providing answers to multiple major aims, as like skills of the laboratories, the operators training, but also the harmonisation of the practices. However, the different inter-laboratory comparisons that exist to this day should be improved to meet the expectations of applicants pursuant.

Key words (thematic and geographical area) :

Inter-laboratory test (ILT) - Proficiency test - Hydrobiology - Collaborative trial - Operator - Laboratory - Sampling - Analysis

1. Introduction

Dans un objectif général d'amélioration de la qualité des données issues des mesures réalisées dans le cadre des programmes de surveillance DCE, des essais interlaboratoires (EIL) sont préconisés dans différents domaines (chimie, hydrobiologie, microbiologie, etc.). Dans ce contexte DCE, l'accréditation des laboratoires implique la réalisation de comparaisons interlaboratoires et particulièrement d'essais d'aptitude. La nécessité d'effectuer ce type d'essais pour les laboratoires est réaffirmée dans l'Arrêté « Agrément des laboratoires » révisé d'octobre 2011 qui précise que pour être agréés, les laboratoires doivent « *participer, au moins une fois par période d'agrément, à des programmes de comparaisons interlaboratoires ou d'essais d'aptitude réalisés en France ou dans un autre Etat membre de l'Union européenne [...] s'ils existent* ». Les méthodes hydrobiologiques concernées par cet agrément sont fixées par l'avis du 4 février 2012.

En hydrobiologie, l'offre en EIL par des OCIL (organismes de comparaison interlaboratoires) est peu importante et était même nulle jusqu'en 2010. La mise en place de ce cadre réglementaire ne peut donc se faire que graduellement et en relation avec l'évolution de l'offre. Ainsi, il est apparu nécessaire de mieux définir ces EIL devenus obligatoires notamment en ce qui concerne leur contenu et l'analyse des résultats obtenus. Cette réflexion a démarré en 2011 notamment lors du séminaire Aquaref sur les incertitudes (décembre 2011). Une présentation de la réflexion globale des EIL en hydrobiologie a également été réalisée lors du séminaire Aquaref de 2013.

Ce rapport propose dans un premier temps de faire le point sur la terminologie employée et les notions d'essais interlaboratoires qu'elle recouvre, et de les rapporter au domaine particulier de l'hydrobiologie. Ensuite, les spécificités des méthodes hydrobiologiques sont rappelées afin de comprendre comment ces EIL, déjà largement organisés dans le domaine de la chimie par exemple, pourraient s'appliquer en hydrobiologie. Dans un troisième temps, une description du contenu général d'un essai EIL en hydrobiologie ainsi que des méthodes permettant l'analyse des résultats est faite. Enfin, un tour d'horizon des EIL existant en Europe est réalisé permettant de répertorier l'offre actuelle en EIL susceptible de répondre aux préconisations de l'Arrêté « Agrément des laboratoires ».

La réflexion sur le contenu et l'analyse des résultats d'essais interlaboratoires en hydrobiologie est focalisée sur les méthodes hydrobiologiques en eaux douces, et nécessitera d'être étendue aux eaux littorales. Plusieurs démarches d'EIL déjà réalisées en milieu marin sont cependant analysées dans le chapitre 6 de ce rapport.

Récemment, une norme européenne encadrant les essais comparatifs pour les mesures hydrobiologiques a été publiée : NF-EN 16101 - Guide pour les études comparatives interlaboratoires ayant pour objet l'évaluation écologique - Décembre 2012. Ses préconisations sont explicitées et prises en compte dans le présent rapport.

Cette réflexion sur la conception des essais interlaboratoires en hydrobiologie a bénéficié d'une collaboration avec Béatrice Lalère, Sophie Lardy-Fontan et Nathalie Guigues, du LNE. Nous les en remercions.

2. Qu'est ce qu'un Essai Interlaboratoires ?

Les comparaisons interlaboratoires ou essais interlaboratoires (CIL ou EIL) sont définis au sein de plusieurs normes européennes comme l'organisation, la réalisation et l'évaluation d'essais ou de mesures sur des objets identiques ou similaires par au moins deux laboratoires différents dans des conditions prédéterminées (ISO/CEI 17043 - 2010, NF ISO 13528 - 2005, NF EN 16101 - 2012, etc.).

Les CIL ou EIL peuvent avoir trois principaux objectifs comme l'illustre la figure 1 ci-dessous.

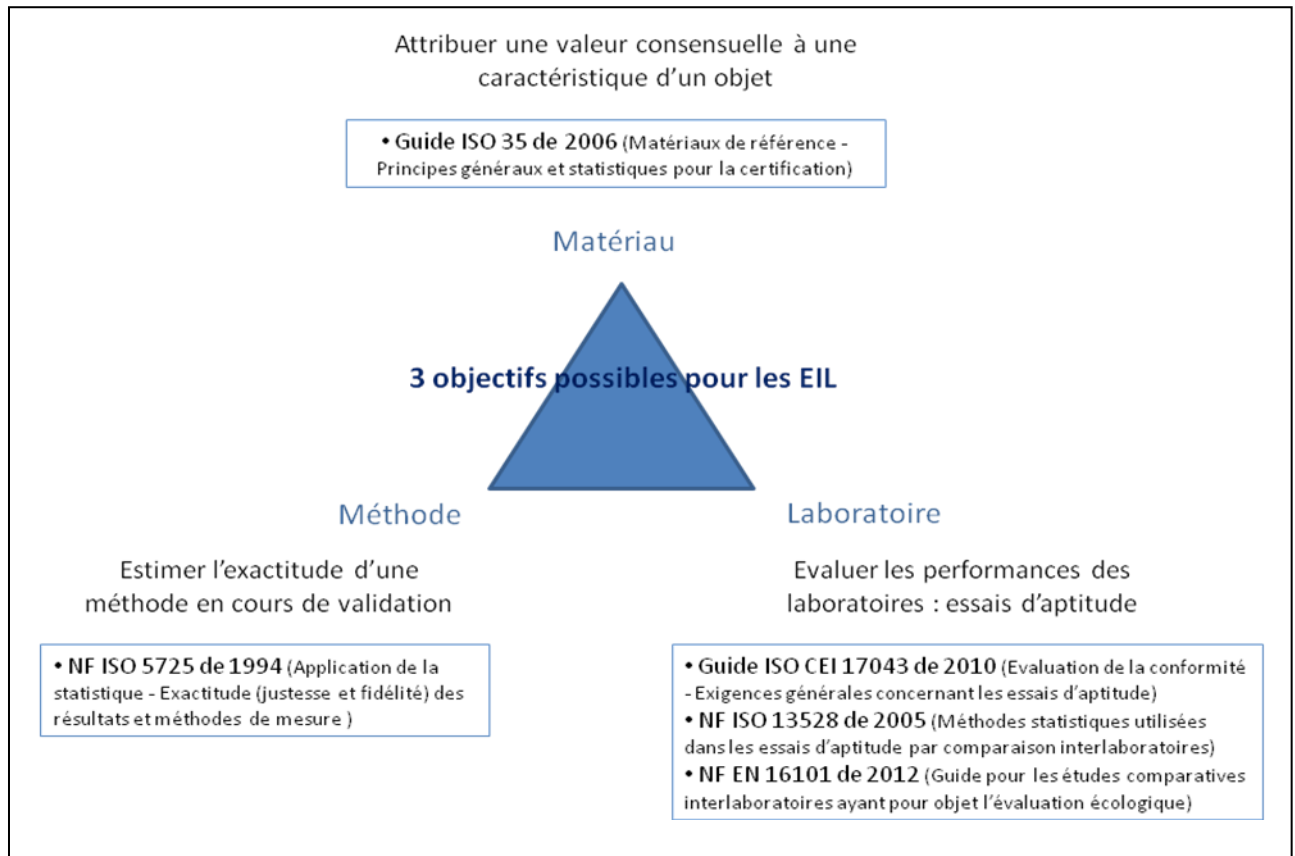


Figure 1 : Objectifs des CIL ou EIL - source LNE.

En hydrobiologie, afin que les laboratoires puissent démontrer leurs compétences et répondre aux exigences des textes d'application de la DCE, c'est-à-dire maîtriser et garantir la qualité des données de surveillance, deux de ces objectifs sont concernés.

- Des comparaisons interlaboratoires peuvent être organisées afin d'évaluer les méthodes hydrobiologiques (estimation de la performance des méthodes, évaluation de la fidélité, de la reproductibilité de ces méthodes par différents laboratoires). Ce type de CIL est par exemple organisé lors d'études sur l'estimation des différentes sources d'incertitudes de mesure par plusieurs laboratoires, telles que les variabilités « opérateurs » ou temporelles (Dagens et Chauvin, 2013). En hydrobiologie, ces approches sont encore réalisées dans le cadre

de programmes de recherche et de développement des méthodes. L'enjeu est ainsi différent des applications en analyse chimique pour lesquelles les essais sont couramment réalisés pour évaluer les performances des méthodes utilisées par chaque laboratoire et le positionnement des résultats obtenus par les laboratoires participants. En hydrobiologie cette application n'a pas la même acception, puisque la méthode utilisée est indissociable de l'élément biologique mesuré, et préconisé de façon réglementaire, tout au moins pour ce qui est des programmes de surveillance dans les réseaux DCE. La norme NF ISO 5725 de 1994 précise les méthodes statistiques devant être utilisées pour ce type d'essais.

- Le deuxième objectif potentiel d'une CIL en hydrobiologie est l'évaluation des performances des laboratoires, appelé aussi **essai d'aptitude**. L'essai d'aptitude en hydrobiologie permet la vérification de l'application des protocoles et l'évaluation des compétences des laboratoires les uns par rapport aux autres. Sa définition exacte en ce qui concerne l'évaluation écologique selon la norme NF 16101 est l' « évaluation de la performance d'un participant selon des critères préétablis, au moyen de comparaisons interlaboratoires ». Parmi les trois normes donnant des indications pour la réalisation des essais d'aptitude la norme NF 16101 est la seule spécifique à l'évaluation écologique.

En dehors de ces différents types d'EIL formels, il existe une autre méthode de comparaison entre laboratoires en hydrobiologie qui a plutôt pour objectif l'échange, la formation ou encore le perfectionnement d'un laboratoire vis-à-vis d'une méthode hydrobiologique : les **essais collaboratifs**. Ces tests ne sont pas destinés à évaluer directement les performances des laboratoires en les positionnant les uns par rapport aux autres mais à permettre à plusieurs laboratoires de comparer leurs résultats afin d'expliquer les différences obtenues et de discuter des difficultés rencontrées. Ces essais collaboratifs ne sont pas définis dans les textes réglementaires. Un essai de ce type est organisé par exemple depuis plus de 10 ans entre les laboratoires des DREAL pour la méthode « diatomées en cours d'eau ». D'autre part, ce type d'approche a déjà été testé dans le domaine de l'échantillonnage de l'eau et des sédiments (Rapport Aquaref 2009 Eau souterraine, par exemple).

Dans ce rapport, seuls les essais interlaboratoires ayant pour objectif premier la comparaison des laboratoires entre eux seront étudiés (essais d'aptitude ou collaboratifs).

3. Spécificité des méthodes hydrobiologiques

3.1. Le protocole d'acquisition de la donnée dans la chaîne d'évaluation

L'objectif global de ces essais est de garantir que, quelque soit le laboratoire qui réalise la mesure, tous les résultats sont comparables et de qualité maîtrisée et suffisante. Il est

donc important que les comparaisons portent sur l'ensemble des phases techniques qui permettent de fournir un résultat.

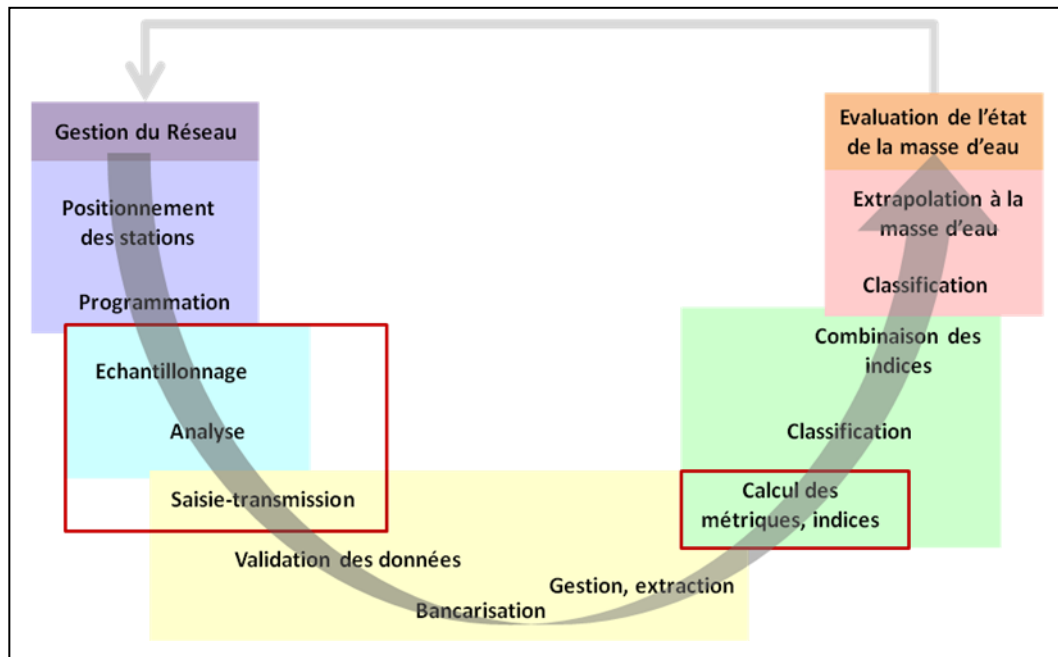


Figure 2 : La chaîne d'évaluation en hydrobiologie.

Les rectangles colorés correspondent à des phases homogènes, les cadres rouges aux étapes impliquant potentiellement le laboratoire d'hydrobiologie.

Dans un réseau de stations de surveillance prédéfini, on peut considérer que le protocole d'acquisition de la donnée hydrobiologique débute avec l'échantillonnage, phase regroupant souvent la délimitation du site selon la méthode préconisée, la description stationnelle, l'élaboration du plan d'échantillonnage et l'échantillonnage lui-même. Selon les méthodes hydrobiologiques, l'échantillonnage peut être un prélèvement pour analyse au laboratoire (les échantillons pouvant être traités ou préparés sur le terrain ou au laboratoire), ou un relevé *in situ*. Lors d'un relevé ou d'un traitement sur place (macrophytes et poissons, par exemple), seuls quelques échantillons à détermination difficile sont prélevés pour validation de la détermination au laboratoire, la plus grande partie de l'identification se faisant sur le terrain. La phase d'analyse peut donc comprendre la préparation et la détermination complète des échantillons prélevés ou uniquement une validation de la détermination des taxons particuliers. Dans l'organigramme d'acquisition des données dans les réseaux DCE, le protocole aboutit à la saisie des données pour transmission aux Agences de l'eau, « responsables des données » au sens du SNDE¹.

Il est important de noter ici que ce qui est défini comme « donnée hydrobiologique » au sens de la DCE et qui est le support à l'application des règles d'évaluation, est le plus généralement constitué d'une liste floristique ou faunistique assortie de l'abondance de chaque taxon.

¹ SNDE : Schéma national des données sur l'eau, mis en place par l'arrêté du 26 juillet 2010

Ce sont donc ces listes elles-mêmes qui constituent le « résultat » brut de la « mesure ». Cependant les laboratoires opérateurs des mesures hydrobiologiques peuvent également avoir à calculer des métadonnées, tels que des indices biologiques, assimilés ou complémentaires aux résultats bruts dans certains cahiers des charges.

3.2. Travailler sur des taxons et des peuplements

3.2.1. Ecorégionalisation

Les méthodes hydrobiologiques, telles qu'elles sont définies pour la plupart des applications d'évaluation d'état des milieux aquatiques (méthodes DCE en particulier), sont basées sur un inventaire faunistique ou floristique. A l'échelle d'un territoire national très varié comme celui de la France, les espèces composant les peuplements ne sont pas les mêmes selon les milieux inventoriés ou les zones biogéographiques dans lesquelles on procède. Si l'on considère par exemple les zones alpines, les plaines de basse altitude ou les régions méditerranéennes, on conçoit que les espèces observées, pour une partie d'entre elles, ne seront pas les mêmes.

Contrairement au domaine de l'hydrochimie, pour lequel la nature des éléments dosés n'est pas liée à des zones géographiques particulières, il existe une écorégionalisation pour les taxons. Par exemple, pour les macrophytes, seulement 21 % des 640 taxons relevés dans les cours d'eau français ont été observés dans au moins 9 hydroécorégions sur les 20 prospectées en France métropolitaine².

Cette caractéristique, nette pour certains groupes taxonomiques à l'échelle de la France continentale, l'est encore plus, bien évidemment, si l'on étend l'application aux départements d'outre-mer, où les peuplements tropicaux n'ont aucun lien avec ceux identifiés en métropole.

La comparaison des résultats obtenus par différents laboratoires, si elle doit être réalisée sur les listes de taxons elles-mêmes, se heurte à cette spécificité potentielle selon les écotypes de cours d'eau. Les comparaisons étant potentiellement applicables au niveau européen, le problème n'en est que plus aigu.

S'agissant de tester les compétences des laboratoires, il faut aussi noter que la plupart des opérateurs œuvrant en hydrobiologie ont des zones d'activité qui ne dépassent pas un ou quelques domaines biogéographiques. On ne pourra donc pas attendre des laboratoires une efficacité équivalente sur des échantillons provenant de l'ensemble des écorégions d'un territoire tel que l'Europe.

3.2.2. Méthodes semi-destructives

La phase d'échantillonnage nécessite généralement des prélèvements sur plusieurs endroits ou des relevés exhaustifs (pêches électriques, par exemple), ne laissant pas les éléments biologiques en place. Ces méthodes peuvent être qualifiées de semi-destructives,

² D'après la base de données Pandore-Irstea, regroupant les données issues des réseaux de surveillance DCE, à la date de novembre 2013.

car elles peuvent impacter les peuplements en place pour une durée de quelques heures à quelques jours.

Cet aspect des approches hydrobiologiques engendre une difficulté dans la mise en œuvre de comparaisons, puisque l'application répétée de tels protocoles sur un site, par un ou plusieurs opérateurs conduit inévitablement à un biais susceptible d'aller jusqu'à l'épuisement total des peuplements en place.

3.3. Phases techniques dans les protocoles hydrobiologiques

L'application d'une méthode hydrobiologique aboutit à une mesure hydrobiologique. Rappelons que la « mesure hydrobiologique », telle que définie dans les applications DCE, fournit une liste taxonomique ainsi que les abondances associées à chaque taxon.

Chaque protocole technique est constitué de plusieurs phases allant du choix de la définition du site à la saisie des données, voire un enchaînement d'opérations définies par des normes distinctes. C'est par exemple le cas de la méthode Invertébrés en cours d'eau pour laquelle une première norme explique l'échantillonnage des macro-invertébrés en cours d'eau et une deuxième norme détaille le traitement et la détermination de l'échantillon obtenu. Ces deux protocoles sont indissociables pour l'application de cette méthode.

Les différentes phases des méthodes hydrobiologiques n'ont pas la même importance potentielle sur la mesure selon les méthodes, comme l'illustre la figure 3 ci-dessous.

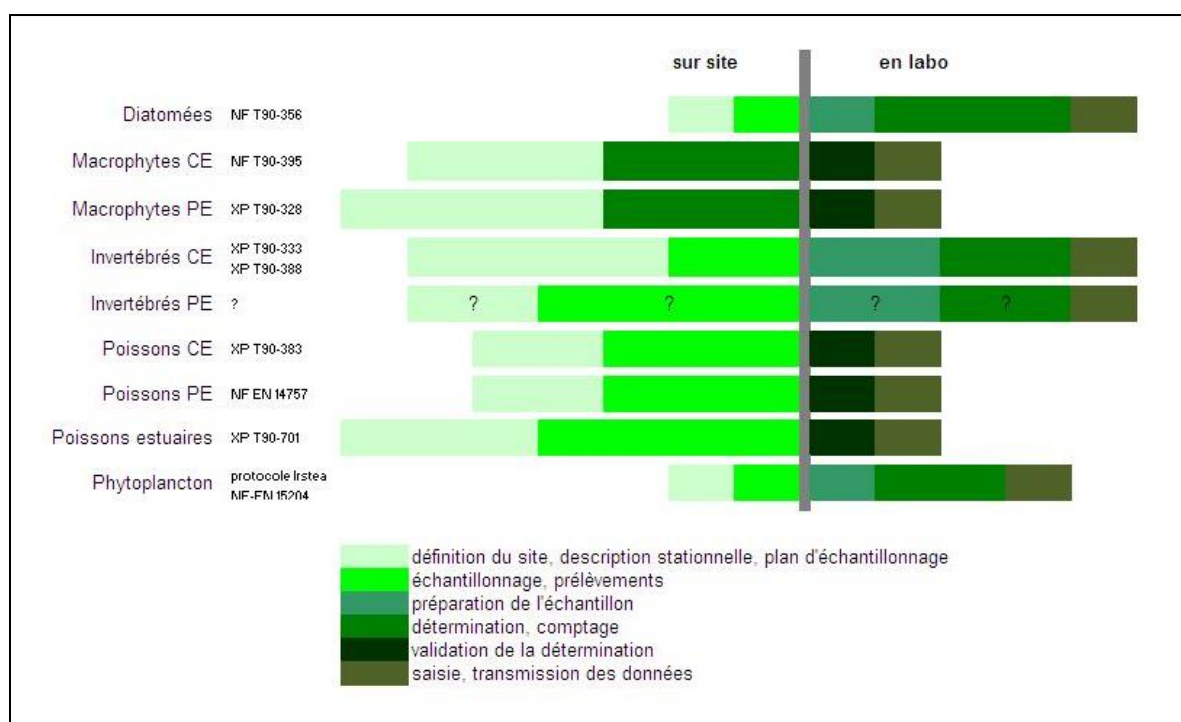


Figure 3 : Importance relative de chaque étape dans les différents protocoles hydrobiologiques utilisés dans les réseaux de surveillance DCE (évaluation par avis d'expert au vu des référentiels techniques).

Selon les méthodes concernées, la phase d'échantillonnage peut être très complexe. Une estimation et une description des habitats de la station est parfois nécessaire pour définir le plan d'échantillonnage, avant de commencer le prélèvement lui-même (invertébrés benthiques et poissons en estuaires, par exemple).

Concernant les incertitudes associées, plusieurs sources d'erreurs peuvent être identifiées, affectables à chacune de ces phases :

- Variabilité spatiale selon le site choisi par l'opérateur ;
- Variabilité intra et inter-opérateurs :
 - o Liée à l'échantillonnage lui-même ;
 - o Liée au traitement de l'échantillon ;
 - o Liée à la détermination ;
- Variabilité temporelle selon la période de réalisation du ou des protocoles ;
- Erreurs de saisie sur le terrain ou au laboratoire.

L'importance relative de toutes ces sources d'incertitude n'est pas encore complètement évaluée. Elle est en cours d'étude pour les méthodes appliquées en réseaux DCE.

Une incertitude potentiellement importante sur le résultat de la mesure a ainsi été estimée sur l'étape d'échantillonnage pour la méthode Macro-invertébrés, engendrant de 7 à 28 % de variabilité sur la note finale de l'indicateur, de part les différences observées entre les opérateurs (Archaimbault 2012).

Certaines méthodes mettent en œuvre des phases « terrain » et « analyse » indissociables. Dans ces cas, la phase d'échantillonnage peut représenter la majeure partie des opérations, l'échantillonnage et la détermination se faisant simultanément sur le terrain (macrophytes, poissons). La partie laboratoire est alors très courte puisque seuls les échantillons à détermination difficile qui auront été prélevés nécessiteront une analyse au laboratoire (validation taxonomique).

Les essais d'aptitude organisés sur ces méthodes hydrobiologiques doivent donc être adaptés à ces spécificités méthodologiques. Leur mise en œuvre est donc significativement différente des tests organisés en routine sur les méthodes d'analyse chimique, mais aussi en fonction des méthodes elles-mêmes, puisque chacune comporte des phases techniques d'importance relative différente.

Outre les écueils liés à l'objet d'étude lui-même, des limites à l'organisation de tels essais d'aptitude peuvent apparaître notamment en ce qui concerne la gestion de l'organisation, le coût ou les possibilités techniques.

4. Contenu d'un essai d'aptitude en hydrobiologie

4.1. Les référentiels applicables

Plusieurs textes, dont certains très récents, constituent une base pour la définition et l'organisation des comparaisons interlaboratoires. Les référentiels recensés sont principalement constitués par 3 normes internationales ou européennes de portée générale et des préconisations spécifiques en annexe d'une norme technique.

4.1.1. NF EN ISO/CEI 17043 d'avril 2010

Evaluation de la conformité - Exigences générales concernant les essais d'aptitude

Cette norme annule et remplace l'ISO/CEI Guide 43-1:1997 et l'ISO/CEI Guide 43-2:1997. Elle s'applique aux comparaisons interlaboratoires destinées à déterminer la performance des laboratoires (essais d'aptitude), ainsi qu'aux activités complémentaires pouvant apporter des preuves indépendantes de la compétence des laboratoires (performance des méthodes par exemple).

Les exigences techniques des essais d'aptitude y sont expliquées notamment en ce qui concerne la conception et l'exécution des programmes d'essai, ainsi que l'analyse des résultats. Des précisions sont également apportées sur les exigences de ces essais relatives au management. Enfin trois annexes informatives apportent des éléments complémentaires sur les différents types de programmes d'essai, les méthodes statistiques pouvant être utilisées lors de l'analyse des données et les principes de sélection et d'utilisation des programmes d'essais d'aptitude par les différentes parties intéressées.

4.1.2. NF ISO 13528 de décembre 2005

Méthodes statistiques utilisées dans les essais d'aptitude par comparaisons interlaboratoires

Cette norme complète le guide ISO / CEI 43 lui-même remplacé par la norme ISO/CEI 17043 évoquée précédemment. Elle détaille les méthodes statistiques devant être utilisées par les organisateurs d'essais d'aptitude pour l'analyse des données mais s'applique uniquement aux données quantitatives.

Cette norme donne les lignes directrices pour la conception et l'interprétation statistiques des essais d'aptitudes. Elle définit et détaille les différentes étapes de l'analyse statistique des résultats quantitatifs des essais d'aptitude et notamment la détermination de la valeur assignée et le calcul des statistiques de performances.

4.1.3. NF EN 16101 de décembre 2012

Guide pour les études comparatives interlaboratoires ayant pour objet l'évaluation écologique

Cette norme européenne a été élaborée spécifiquement en appui à l'application des méthodes hydrobiologiques pour la surveillance DCE, pour répondre à la demande faite au CEN par la Commission européenne de développer les référentiels normatifs utiles à la mise en œuvre des aspects techniques de la DCE. Ce texte est le seul référentiel qui encadre l'application des essais interlaboratoires pour les mesures hydrobiologiques.

Cette norme apporte des informations sur les modes opératoires à appliquer lors de la réalisation de comparaisons interlaboratoires pour les méthodes biologiques en ce qui concerne les travaux sur le terrain et les travaux au laboratoire. Elle précise également quelques méthodes et approches statistiques à considérer pour l'analyse des résultats.

4.1.4. PrEN 14184 (mars 2012, version finale en cours d'approbation)

Guide pour l'étude des macrophytes aquatiques dans les cours d'eau

Cette norme européenne publiée en 2003, dont la révision s'achève, précise désormais les principes de la comparaison interlaboratoires pour les relevés de macrophytes dans les eaux courantes dans son annexe A. Le contenu de cette annexe reprend les discussions ayant eu lieu au CEN (TC230 WG2) concernant la réalisation d'essais interlaboratoires sur les différents éléments biologiques. Une norme cadre générale n'a pas encore été proposée, mais ces éléments, adaptés aux macrophytes en cours d'eau, ont été intégrés à cette norme technique et permettent ainsi de préciser la norme NF EN 16101.

Ce texte explique les différents aspects de la méthode « macrophytes en cours d'eau » devant être intégrés à une CIL. Il précise également les méthodes d'analyse des résultats pouvant être utilisées aussi bien pour les données qualitatives que pour les données quantitatives. Il met en particulier l'accent sur la nécessaire adaptation de l'exploitation des résultats au caractère qualitatif (« catégoriel ») des données brutes, comme les classes de recouvrement pour estimer l'abondance.

4.2. Quelles phases méthodologiques dans un essai d'aptitude ?

L'essai d'aptitude doit permettre de comparer la capacité des laboratoires à réaliser l'acquisition de données selon une méthode précise, avec une bonne fiabilité des résultats. Comme précisé dans la norme NF EN 16101, il est donc nécessaire d'y intégrer toutes les étapes de l'acquisition de la donnée susceptible d'avoir un effet notable sur le résultat final : stratégie d'échantillonnage, modes opératoires du prélèvement, traitement des échantillons, analyse des échantillons, validation des listes taxonomiques.

Pour une grande partie des méthodes hydrobiologiques, la phase d'échantillonnage, actuellement rarement prise en compte dans les essais d'aptitude, ne peut donc pas être dissociée de la phase analytique (préparation des échantillons, tri, détermination, dénombrement).

En principe, toutes les étapes sont donc à prendre en compte dans un essai d'aptitude. Une nuance peut néanmoins être faite pour certaines méthodes hydrobiologiques qui présentent des phases simples et dont l'impact est *a priori* négligeable sur le résultat de la mesure. Si l'effort pour intégrer ces étapes est disproportionné au vu du gain attendu dans la pertinence de la comparaison, ou que ces étapes sont encadrées dans un autre processus, le protocole de l'essai pourra éventuellement les exclure. Ce sera par exemple le cas pour le phytoplancton, pour lequel la phase d'échantillonnage se rattache à un prélèvement d'eau, et peut être contrôlée en tant que telle. Le test visant les compétences hydrobiologiques pourra donc se focaliser sur les aspects de traitement de l'échantillon et de détermination taxonomique/comptage.

L'essai d'aptitude en hydrobiologie vise la comparaison des résultats bruts de « l'analyse ». Il ne doit donc pas *a priori* se baser uniquement sur des métadonnées issues de l'exploitation des résultats, comme par exemple un indice biologique. De plus, les

compétences mises en œuvre pour le calcul d'indice lui-même (utilisation de référentiels, souvent par l'intermédiaire de logiciels dédiés ou de tableurs informatiques) ne sont pas de l'ordre de la production de données hydrobiologiques, mais de l'exploitation de ces données. A ce titre, on peut considérer qu'elles n'entrent pas dans le champ des compétences précises que l'on veut tester.

Toutefois, l'évaluation de qualité ou d'état des milieux aquatiques étant l'objectif le plus courant de ces mesures hydrobiologiques, le calcul de certains indicateurs (métriques, indices) peut être inclus dans les comparaisons, en complément au traitement des résultats bruts. Cela peut renseigner utilement sur l'impact des éventuels écarts constatés dans les résultats bruts sur quelques indicateurs utilisés en exploitation de ces données.

La norme NF EN 16101 propose deux approches différentes concernant le mode opératoire convenant à ces essais d'aptitude.

L'approche progressive permet de réaliser l'essai en plusieurs fois en évaluant les différentes phases de la méthode hydrobiologique séparément :

- Evaluation des compétences d'identification en laboratoire (ensemble d'organismes de composition taxonomique connue préparé par un expert) ou sur le terrain selon les méthodes ;
- Evaluation de la stratégie d'échantillonnage et des données présentées, sur le terrain ;
- Evaluation des résultats suite à l'analyse par les participants en laboratoire.

L'approche alternative consiste en l'élaboration d'un essai d'aptitude en une seule phase comprenant :

- Echantillonnage de sites présélectionnés et détermination de la composition taxonomique et de l'abondance par l'examineur ;
- Echantillonnage de sites présélectionnés et détermination de la composition taxonomique et de l'abondance par les participants ;
- Evaluation des résultats des participants présentés à l'examineur.

4.3. Comparer des analyses semi-destructives

L'impact des prélèvements eux-mêmes sur le site d'échantillonnage dû à l'application de certains types de prélèvements ou relevés peut engendrer d'importantes difficultés dans la réalisation d'un essai d'aptitude, plusieurs échantillonnages par des opérateurs différents devant être réalisés dans un court laps de temps et sur un même site.

Deux options sont proposées dans la norme NF EN 16101 pour pallier aux difficultés rencontrées si l'échantillonnage d'une méthode est destructif :

- Plusieurs participants procèdent à l'échantillonnage sur des sites multiples mais similaires ;
- Des examineurs ou un jury d'experts indépendants préparent des échantillons artificiels multiples devant être analysés par des participants multiples.

Cependant, ces suggestions doivent être adaptées en fonction des méthodes car selon les éléments biologiques de qualité concernés, des échantillons synthétiques ne pourront pas toujours être préparés (poissons, macrophytes).

D'autre part, la notion de « sites similaires » est délicate à interpréter. A l'échelle du prélèvement, c'est-à-dire l'habitat ou le micro-habitat, la variabilité spatiale d'un cours d'eau ou, *a fortiori*, de plusieurs cours d'eau, rend a priori peu pertinente la comparaison fine des résultats obtenus sur des stations différentes. Il paraît difficile de maîtriser l'influence de l'inter-station dans la vision inter-opérateurs.

Une autre possibilité pourrait consister en l'évaluation de chaque participant à l'essai d'aptitude par rapport aux résultats obtenus par un laboratoire de référence, lors d'essais qui ne mettent en jeu à chaque session que le laboratoire de référence et un seul laboratoire testé, afin de limiter la pression sur le site à seulement deux intervenants.

C'est ici la notion de « laboratoire de référence » qui est difficile à interpréter, puisqu'il n'est pas possible de garantir l'exhaustivité d'un prélèvement ou d'un dénombrement, même si les compétences de l'opérateur sont reconnues comme très élevées. La méthode utilisée par le référent pourrait alors être différente (effort d'échantillonnage supérieur, répétitions de la préparation de l'échantillon et des déterminations), pour mieux correspondre à la notion de référence. Mais c'est alors un biais correspondant à la différence de méthode utilisée qui risquerait d'interférer.

Pour certains éléments biologiques, même la limitation à seulement deux opérateurs successifs reste de toute manière destructive (pêche électrique, en particulier).

4.4. Problématique de régionalisation

La portée géographique d'un essai d'aptitude étant au minimum nationale (cas des démarches d'accréditation), les différents laboratoires participants sont amenés à réaliser cet essai dans une écorégion avec laquelle ils ne sont pas tous très familiers. Les espèces présentes dans les échantillons de l'exercice et celles effectivement rencontrées par les opérateurs dans les sites qu'ils étudient peuvent donc différer, principalement dans les habitats caractéristiques des types de cours d'eau considérés. De ce fait, un biais peut être attendu dans l'évaluation de la capacité réelle des opérateurs dans leurs mesures de routine, ce qui est le but recherché par ces tests.

Dans ce contexte, les essais collaboratifs semblent être mieux adaptés, surtout s'ils sont organisés dans des écorégions différentes à chaque session, comme c'est le cas par exemple des essais annuels menés par les laboratoires de DREAL sur les diatomées.

A l'échelle européenne, ce biais est *a priori* accentué. Toutefois, pour certains organismes à répartition paneuropéenne, ces tests sont envisageables (ring-test « phytoplancton », en particulier).

Lorsqu'un essai est basé sur l'examen d'échantillons synthétiques, il peut être réalisé un échantillon de référence combinant plusieurs échantillons provenant des différents secteurs de chaque participant (échantillons fournis par les participants). Cet échantillon peut ainsi comprendre des spécimens de différentes régions. Ce type d'essai ne permet pas de représenter complètement la méthode d'évaluation et s'éloigne de l'application en

routine mais permet d'évaluer la capacité de détermination de taxons particuliers à l'échelle nationale.

Pour pallier ce genre de difficultés, il peut être possible de réaliser les essais d'aptitude dans des régions différentes chaque année, sur un cycle de plusieurs années. Ce système cyclique pourrait assurer un maintien des compétences dans le temps avec un panel plus large de situations régionales testées. De plus, un tel système aurait l'avantage de développer les compétences en permettant aux opérateurs de se confronter à des taxons ou des habitats dont ils n'ont pas la pratique. Il est cependant à noter que ce changement de régions d'une année sur l'autre peut entraîner une plus forte hétérogénéité dans les échantillonnages. En effet, la structure des cours d'eau selon les régions pouvant être très différente, la variabilité sur le terrain sera probablement plus importante si les essais sont faits chaque année dans une région différente que si les essais sont toujours réalisés au même endroit.

4.5. Conditions de réalisation

Il est important que l'essai d'aptitude reflète au maximum les conditions réelles dans lesquelles les protocoles sont appliqués en routine par chaque laboratoire, afin de positionner les résultats dans le contexte où les données de surveillance sont effectivement produites.

Certains paramètres dépendant des pratiques ou des procédures internes de chaque laboratoire devrait être considérées. Par exemple :

- Méthodes de conservation des échantillons : selon les fixateurs, les techniques utilisables pour le comptage et le dénombrement ne sont pas toujours les mêmes (cas de la congélation des échantillons d'invertébrés, par exemple) ;
- Informations sur l'origine des échantillons fournis pour la réalisation des essais : la non-connaissance de cette donnée peut avoir une influence sur les résultats car elle peut être nécessaire lors de la détermination taxonomique ;
- Type de matériel et mise en œuvre lors de l'échantillonnage.

4.6. Adapter les essais à chaque méthode

Comme précisé dans la norme NF EN 16101, « pour chaque élément biologique, l'approche méthodologique concrète de la comparaison des analyses de terrain sera différente ». En effet, les différentes étapes du protocole n'ont pas toutes le même poids selon la méthode hydrobiologique concernée. Par exemple, les opérations de terrain sont prépondérantes par rapport aux travaux de laboratoire pour certaines méthodes, alors qu'elles ont une importance relative moindre pour d'autres. On peut donc admettre que l'incertitude de chaque phase méthodologique n'a pas un impact équivalent sur le résultat final. Le contenu des essais doit donc être adapté à chaque protocole.

Les tableaux ci-contre (figure 4) proposent une hiérarchisation des phases qui devraient être intégrées aux essais d'aptitude, pour chacun des protocoles utilisés actuellement dans les programmes de surveillance de l'état écologique des milieux aquatiques.

Figure 4 : Intégration des différentes phases de l'acquisition de la donnée aux essais d'aptitude selon les méthodes hydrobiologiques (éléments biologiques et méthodes préconisées d'après l'arrêté Surveillance du 29 juillet 2011 et les accréditations existantes).

Légende des tableaux :

	Méthodes concernées par l'Arrêté Agrément : publiées dans l'avis du 4 février 2012
○	Sans objet
●	Possibilité de ne pas intégrer l'étape à l'EIL car peu d'impact sur la donnée
●	Étape à intégrer obligatoirement à l'EIL
(●)	Étape complémentaire non obligatoire dans le protocole d'acquisition de la donnée mais à intégrer obligatoirement à l'EIL

Catégories de Masse d'eau	Intitulé "arrêté Surveillance"	Intitulé "support" SANDRE	Méthode - Document de référence technique	Phases des méthodes d'acquisition de données						
				Définition du site, description stationnelle, plan d'échantillonnage	Echantillonnage, prélèvement	Traitement de l'échantillon	Détermination, comptage terrain	Détermination, comptage laboratoire	Validation du résultat de la mesure hydrobiologique	
COURS D'EAU	Faune benthique invertébrée	Macro-invertébrés benthiques	NF T90-350 (2004-03-01). Qualité de l'eau - Détermination de l'indice biologique global normalisé (IBGN).	●	●	●	○	●	●	
			XP T90-333 (2009-09-01) : Prélèvement des macroinvertébrés aquatiques en rivières peu profondes.	●	●	●	○	●	●	
			XP T90-388 (2010-06-01) : Qualité de l'eau - Traitement au laboratoire d'échantillons contenant des macro-invertébrés de cours d'eau.	●	●	●	○	●	●	
		Protocole expérimental d'échantillonnage des macroinvertébrés en cours d'eau profonds. P. Usseglio-Polatera, JG. Wasson, V. Archaimbault. Note technique Université de Metz - Cemagref, décembre 2009.	●	●	●	○	●	●		
		XP T90-388 (2010-06-01) : Qualité de l'eau - Traitement au laboratoire d'échantillons contenant des macro-invertébrés de cours d'eau.	●	●	●	○	●	●		
		Oligochètes	NF T90-390 (2002-04-01). Qualité de l'eau - Détermination de l'indice oligochètes de bioindication des sédiments (IOBS).	●	●	●	○	●	●	
		Phytobenthos	Diatomées benthiques	NF T90-354 (2007-12-01). Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD).	●	●	●	○	●	●
		Macrophytes	Macrophytes	F T90-395 (2003-10-01). Qualité de l'eau - Détermination de l'indice biologique macrophytique en rivière (IBMR).	●	○	○	●	●	●
		Phytoplancton	Phytoplancton	Protocole standardisé d'échantillonnage et de conservation du phytoplancton en grands cours d'eau applicable aux réseaux de mesure DCE. C. Laplace-Treyture, C. Chauvin, M. Menay, A. Dutartre. Rapport Cemagref, Version 2. Décembre 2010.	●	●	●	○	●	●
	NF EN 15204 (2006-12-01). Qualité de l'eau - Norme guide pour le dénombrement du phytoplancton par microscopie inversée (méthode Utermöhl).			●	●	●	○	●	●	
	Ichtyofaune	Poissons	XP T90-383 (2008-05-01). Qualité de l'eau - Échantillonnage des poissons à l'électricité dans le cadre des réseaux de suivi des peuplements de poissons en lien avec la qualité des cours d'eau.	●	●	○	●	(●)	●	
			NF T90-344 (2004-05-01). Qualité de l'eau - Détermination de l'indice poissons rivière (IPR).	●	●	○	●	(●)	●	

Catégories de Masse d'eau	Intitulé "arrêté Surveillance"	Intitulé "support" SANDRE	Méthode - Document de référence technique	Phases des méthodes d'acquisition de données					
				Définition du site, description stationnelle, plan d'échantillonnage	Echantillonnage, prélèvement	Traitement de l'échantillon	Détermination, comptage terrain	Détermination, comptage laboratoire	Validation du résultat de la mesure hydrobiologique
PLANS D'EAU	Faune benthique invertébrée	Oligochètes	NF T90-391 (2005-03-01). Qualité de l'eau - Détermination de l'indice oligochètes de bioindication lacustre (IOBL).	●	●	●	○	●	●
		Macro-invertébrés benthiques	Verneaux, V., Verneaux J., Schmitt A., Loyv C., Lambert JC., 2004. Indice Biologique Lacustre (IBL) - The Lake Biotic Index (LBI) : an applied method for assessing the biological quality of lakes using macrobenthos ; the Lake Châlain (French Jura) as an example. Ann. Limnol. - Int. J. Lim. 2004, 40 (1), 1-9.	●	●	●	○	●	●
			Mouthon, J., 1993. Indice Mollusque (IMOL) - Un indice biologique lacustre basé sur l'examen des peuplements de mollusques. - Bull. Fr. Pêche Piscic., 1993, 331 : 397-406.	●	●	●	○	●	●
			Protocole d'échantillonnage des invertébrés benthiques adapté aux plans d'eau naturels profonds. L. Mazzella, J. de Bortoli, C. Argillier. Note technique Cemagref, sept 2009.	●	●	●	○	●	●
	Macrophytes	Macrophytes	XP T 90-328 (2011-01-18). Echantillonnage des communautés de macrophytes en plans d'eau.	●	○	○	●	●	●
	Phytoplancton	Phytoplancton	Protocole standardisé d'échantillonnage, de conservation, d'observation et de dénombrement du phytoplancton en plan d'eau pour la mise en oeuvre de la DCE. Version 3.3.1. C. Laplace-Treytore, J. Barbe, A. Dutartre, JC. Druart, F. Rimet, O. Anneville. Rapport Cemagref - INRA. Septembre 2009.	●	●	●	○	●	●
			NF EN 15204 (2006-12-01). Qualité de l'eau - Norme guide pour le dénombrement du phytoplancton par microscopie inversée (méthode Utermöhl).	●	●	●	○	●	●
	Ichtyofaune	Poissons	NF EN 14757 (2005-11-01). Qualité de l'eau - Échantillonnage des poissons à l'aide de filets maillants.	●	●	○	●	(●)	●
EAUX DE TRANSITION	Ichtyofaune	Poissons	XP T90-701 (2011-06-01) - Qualité de l'eau - Échantillonnage au chalut à perche des communautés de poissons dans les estuaires.	●	●	○	●	(●)	●
EAUX LITTORALES	EAUX COTIERES ET AUTRES EAUX DE TRANSITION		NON ENCORE INTEGREES A LA REFLEXION SUR LE CONTENU ET L'ANALYSE						

Dans ces tableaux, les informations sont données à titre indicatif et, dans l'état actuel de la réflexion et de la validation de l'ensemble des acteurs impliqués et des donneurs d'ordre, ne constituent pas un référentiel technique sur le contenu des EIL.

Certaines méthodes ne sont encore pas assez définies ni stabilisées pour envisager concrètement des essais interlaboratoires. C'est par exemple le cas pour les protocoles concernant les invertébrés en plans d'eau.

On notera que des préconisations concernant les CIL commencent à être intégrées directement dans les référentiels techniques, comme par exemple la nouvelle version de la norme européenne « étude des macrophytes en cours d'eau » (prEN 14184), dont une annexe précise les principes de réalisation d'un EIL pour cette méthode.

Dans certains cas, plusieurs protocoles d'échantillonnage existent (selon les milieux considérés), encadrant la collecte d'échantillons qui sont ensuite traités selon un protocole commun (invertébrés en petits ou en grands cours d'eau, en particulier). Dans ce type de cas, il est envisageable de prévoir un essai collaboratif pour chaque protocole d'échantillonnage, complété par un essai commun pour les phases menées en laboratoire. On peut aussi imaginer deux types d'essais regroupant terrain et laboratoire, chacun adapté à un milieu spécifique. Une partie du contenu de ces essais serait alors identique (travaux en laboratoire).

Il faut en effet prendre en compte le fait que les laboratoires peuvent n'opérer en routine que sur un type de milieux (cours d'eau prospectables à pieds, par exemple), car les moyens matériels et humains peuvent être significativement différents selon les protocoles d'échantillonnage et leurs adaptations aux grands cours d'eau. Le but recherché étant toujours, bien évidemment, de dimensionner de façon pertinente le test pour estimer le niveau de maîtrise de la qualité finale des données effectivement produites en routine.

En ce qui concerne la validation du résultat de la mesure hydrobiologique, la nécessité de son intégration aux essais d'aptitude peut être discutée. Dans le cas où cette étape n'est pas intégrée à l'essai, il peut être considéré que l'évaluation de la performance omet une partie importante de l'acquisition de la donnée finale. Cependant, un biais existe également si cette étape est intégrée à l'essai puisque certains opérateurs transmettent leurs échantillons et résultats à des experts pour effectuer cette validation. Dans ce cas, l'essai d'aptitude ne permettra pas réellement l'évaluation des performances de ces laboratoires eux-mêmes.

5. Analyse des résultats

La mesure hydrobiologique est constituée d'une liste de taxons dont l'élaboration est basée sur les capacités de l'opérateur à reconnaître ces taxons, puis des abondances de présence associées à chacun des taxons (cf. norme NF EN 16101). L'analyse des résultats des essais d'aptitude en hydrobiologie doit viser principalement les aspects à la fois qualitatifs (identification des taxons) et quantitatifs (abondances : nombre d'individus ou pourcentage de recouvrement).

Bien que des métriques et indices soient calculés à partir des listes d'espèces afin d'aboutir à l'évaluation écologique et qu'il puisse être informatif d'intégrer ces calculs à un essai d'aptitude, il n'est pas suffisant de concentrer l'analyse des résultats d'un tel essai sur ces valeurs d'indices. En effet, le calcul de ces métadonnées ne rend compte que de certains

aspects des résultats : il est possible d'obtenir des valeurs d'indices et des classes de qualité similaires même si des taxons ont été omis ou mal identifiés.

L'évaluation des compétences des laboratoires uniquement sur la production d'une valeur d'indice ou, *a fortiori*, d'une classe de qualité ou d'état, ne permet que très partiellement d'évaluer les compétences dans le recueil de la donnée de base.

De plus, rappelons que les indices calculés actuellement dans les règles d'évaluation ont vocation à évoluer. Des essais interlaboratoires robustes doivent donc permettre de montrer avant tout la capacité à produire des données de base fiables, utilisables quels soient le mode d'exploitation ou les règles d'évaluation appliquées à ces données.

Les méthodes pouvant être utilisées pour analyser les données des essais d'aptitude sont décrites dans la norme NF ISO 13528 et dans l'annexe B de la norme NF EN ISO/CEI 17043 pour tous les types d'essais d'aptitude. Ces méthodes sont également précisées par Stribling *et al.* (2003) en ce qui concerne les données taxonomiques et dans la norme NF ISO 16101 pour les essais d'aptitude ayant pour objet l'évaluation écologique.

5.1. Données et méthode

Le traitement des résultats de ce type d'essai ne peut pas se faire uniquement à l'aide d'une approche statistique. En effet d'après la norme NF EN ISO/CEI 17043, les résultats d'un mesurage quantitatif (valeurs numériques et rapportées sur une échelle linéaire) peuvent être analysés statistiquement, alors que cette approche est plus délicate pour des résultats qualitatifs (données descriptives et reportées sur une échelle nominale ou ordinale).

Les normes NF ISO 13528 et NF EN ISO/CEI 17043 soulignent la nécessité de réaliser plusieurs étapes telles que :

- la détermination de la valeur assignée, la valeur assignée étant la « valeur attribuée à une grandeur particulière et reconnue, parfois par convention, comme ayant une incertitude appropriée à un usage donné » ;
- le calcul des statistiques de performances : transformation des résultats d'un essai d'aptitude en statistique de performances qui permet de faciliter l'analyse ;
- l'évaluation de la performance qui correspond à la phase d'interprétation et de comparaison des données des essais d'aptitude.

La norme NF ISO 16101 fait quelques recommandations concernant l'analyse statistique des résultats d'essais en écologie. Il est tout d'abord important de s'assurer que les données suivent une distribution normale pour appliquer les méthodes statistiques classiques. Ensuite, des difficultés pour effectuer des comparaisons statistiques peuvent exister si le nombre d'observations disponibles est faible et dans ce cas, des comparaisons graphiques pourront être préférables à des tests statistiques. Aussi, l'analyse statistique multivariée peut être utilisée si plusieurs caractéristiques ont été observées simultanément sur les mêmes objets. Cependant, ce type d'analyse nécessitant des données proches de la distribution normale multivariée (hypothèse généralement pas satisfaite), il sera souvent préférable d'utiliser des comparaisons graphiques. De plus, il sera nécessaire de vérifier que l'analyse des données multivariées n'est pas contaminée par des valeurs aberrantes.

5.2. Détermination de la valeur assignée : NF ISO 13528 et NF EN ISO/CEI 17043

D'après ces deux normes, en ce qui concerne les données quantitatives, la valeur assignée peut être déterminée de différentes façons :

- Valeurs connues (résultats déterminés par une formulation spécifique de l'entité soumise à l'essai d'aptitude) ;
- Valeurs de référence certifiées (déterminées par des méthodes d'essai ou de mesures définitives) ;
- Valeurs de référence (déterminées par analyse, mesurage ou comparaison de l'entité soumise à l'essai d'aptitude avec un matériau de référence ou un étalon) ;
- Valeurs consensuelles provenant des résultats des participants experts ;
- Valeurs consensuelles provenant des résultats des participants.

En hydrobiologie, la « valeur connue » ne peut pas être définie, puisqu'il n'est pas possible de connaître avec certitude les espèces composant un peuplement naturel et leur abondance. Le seul cas où une telle valeur pourrait être définie est celui des échantillons synthétiques, pour les éléments biologiques pour lesquels cela est envisageable (phytoplancton, par exemple). Toutefois, même dans ces cas d'échantillons issus de cultures homogénéisées, la valeur assignée formelle peut difficilement être rigoureusement garantie, en particulier pour les abondances (comptage).

Pour les valeurs qualitatives (ou semi-quantitatives), la valeur assignée est en général déterminée par le jugement d'un expert ou par fabrication (valeur de référence). Une valeur consensuelle peut également être définie par un accord sur un pourcentage majoritaire prédéterminé de réponses (moyenne robuste des valeurs obtenues par les laboratoires).

5.3. Calcul des statistiques de performances

Sont synthétisées ici les préconisations des différents référentiels applicables et autres documents scientifiques répertoriés.

5.3.1. NF ISO 13528 et NF EN ISO/CEI 17043

Ces deux normes suggèrent l'utilisation de différents calculs pour analyser les données quantitatives d'un essai d'aptitude.

D : différence

$$D = (x - X)$$

x : résultat du participant

X : valeur assignée

D% : différence en pourcentage

$$D_{\%} = \frac{(x - X)}{X} \times 100$$

Z-score

$$z = \frac{x - X}{\sigma}$$

L'écart-type σ représente la variabilité des résultats des différents laboratoires.

Il peut être calculé à partir :

- D'une aptitude à l'objectif de performances visé, déterminé par le jugement d'un expert ou un mandat réglementaire (méthode par prescription);
- D'une estimation à partir des cycles d'essais d'aptitudes précédents ou de prévisions fondées sur l'expérience (méthode par perception);
- D'une estimation à partir d'un modèle statistique (modèle général);
- Des résultats d'une expérience en matière de fidélité;
- Des résultats d'un participant, c'est-à-dire un écart-type classique ou robuste fondé sur les résultats des participants.

Le z-score représente un écart par rapport à une valeur de référence. Sa définition dans la norme NF ISO 13528 est une « mesure normalisée du biais de laboratoire, calculée à partir de la valeur assignée et de l'écart-type pour l'évaluation de l'aptitude ».

Zéta-score

$$\zeta = \frac{x - X}{\sqrt{u_{lab}^2 + u_{av}^2}}$$

u_{lab} : incertitude-type composée des résultats d'un participant;

u_{av} : incertitude type de la valeur assignée.

Nombres En : Ecart normalisés

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

U_{lab} : incertitude élargie des résultats d'un participant;

U_{ref} : incertitude élargie de la valeur assignée du laboratoire de référence.

Ces calculs pourraient *a priori* être appliqués sur des données hydrobiologiques quantitatives telles que l'abondance.

D'après ces deux normes, la méthode qui convient pour analyser les résultats d'un essai d'aptitude amenant à des **données qualitatives** est une comparaison du résultat d'un participant à la valeur assignée. Si les deux valeurs sont identiques, la performance sera acceptable. En revanche, si les deux valeurs sont différentes, le jugement d'un expert sera nécessaire pour déterminer la validité du résultat.

Le calcul du Z-score n'est pas approprié pour l'analyse des résultats qualitatifs. En revanche, il est également possible de lister la distribution des résultats des participants avec le nombre ou le pourcentage des résultats dans chaque catégorie et de fournir un résumé des mesures. Il est aussi possible d'évaluer les résultats comme étant acceptables selon la proximité avec la valeur assignée ou d'évaluer les performances en se basant sur les percentiles (ex : les 5 % de résultats les plus éloignés du mode ou de la valeur assignée sont inacceptables).

5.3.2. Stribling et al. (2003)

Dans leurs travaux concernant la qualité des données taxonomiques (qualitatives et quantitatives), Stribling *et al.* suggèrent l'utilisation d'autres calculs que ceux présentés ci-dessus pour l'analyse des résultats taxonomiques des essais d'aptitude.

Exactitude : proximité entre une mesure et une valeur vraie ou vérité analytique

Fidélité : proximité entre différentes mesures de même type

En taxonomie, il faut comparer les résultats d'analyse d'un échantillon sélectionné aléatoirement analysé par 2 taxonomistes ou 2 laboratoires.

La fidélité peut être quantifiée pour l'identification et pour le dénombrement.

- Pourcentage de différences dans le dénombrement (PDE) : détermine la fidélité de comptage (**données quantitatives**)

$$PDE = \frac{|n1 - n2|}{n1 + n2} \times 100$$

n1 = nombre d'individus comptés dans l'échantillon par le 1^{er} participant

n2 = nombre d'individus comptés dans l'échantillon par le 2^{ème} participant

- Pourcentage de désaccords taxonomiques (PDT) : compare les résultats taxonomiques entre 2 participants (**données qualitatives**)

$$PTD = \left[1 - \left(\frac{comp_{pos}}{N} \right) \right] \times 100$$

Comp_{pos} = nombre de comparaisons positives

N = nombre total d'individus sur les 2 comptages

Plus la valeur de PTD est faible, plus la fidélité taxonomique est importante

Il semble que ces formules puissent être utilisées sur des résultats hydrobiologiques, en appliquant le PDE en ce qui concerne les données de comptage (abondance) et le PDT pour comparer les listes taxonomiques obtenues.

5.3.3. Définition des incertitudes - Protocole Invertébrés en cours d'eau

Plusieurs travaux sont actuellement engagés, visant à identifier et quantifier les sources d'incertitudes inhérentes à la chaîne d'opération menant à l'évaluation de l'état écologique des masses d'eau.

Le protocole préconisé dans la surveillance DCE pour l'élément « Invertébrés en cours d'eau » (dénommé Protocole RCS de 2006 à 2010), actuellement transcrit en référentiel technique par les normes XP-T90-333 et XP-T90-388, a fait l'objet d'un programme de recherche mené par Irstea (Archambault, 2012) en collaboration avec les DREAL et plusieurs bureaux d'études prestataires des réseaux. Les phases d'échantillonnage et de traitement en laboratoire ont été analysées par comparaison entre les pratiques et les résultats de plusieurs opérateurs. Les enseignements qui en ont été tirés rejoignent la problématique des EIL, puisque la démarche a été celle d'une intercomparaison sur un certain nombre de critères et de métriques.

De plus, l'analyse des incertitudes apporte des précisions essentielles sur les phases susceptibles d'induire le plus de variabilité entre opérateurs, donc sur celles sur lesquelles des EIL de routine devraient se focaliser plus particulièrement.

Bien que cette étude ait été faite dans le cadre d'un EIL sur la méthode et non sur les laboratoires (cf chap. 1), la comparaison des résultats obtenus par différents opérateurs s'avère tout à fait utilisable pour un essai d'aptitude, notamment en ce qui concerne les données qualitatives (listes de taxons).

Les résultats de cette étude ont été analysés en calculant, entre autres les critères suivants :

- les pourcentages de taxons communs entre les listes des participants et une liste de référence ;
- les pourcentages de taxons trouvés en plus dans la liste de référence ;
- les pourcentages de taxons non retrouvés dans la liste de référence.

Ces travaux montrent l'utilisation de métriques relativement simples pour analyser les données qualitatives des différentes autres méthodes hydrobiologiques dans le cadre d'un essai d'aptitude.

5.3.4. Autres types de calculs

La norme XP CEN ISO/TS 22117 traite de la comparaison de performances des laboratoires pour la microbiologie des aliments. Les méthodes d'évaluation pour les **données quantitatives** sont proches de celles exposées dans les documents cités précédemment. En revanche, les méthodes d'évaluation pour les **données qualitatives** montrent des possibilités différentes de calcul pour

l'analyse de résultats présentés sous la forme « détecté » ou « non détecté » pouvant être comparables aux résultats obtenus en hydrobiologie (« taxon observé » ou « taxon non observé ») :

- Nombre de résultats positifs et négatifs (présence/absence)
- Taux de spécificité

$$r_{SP} = \frac{n_-}{E(n_{-tot})} \times 100\%$$

n_- : nombre de résultats négatifs trouvés

$E(n_{-tot})$: nombre total d'échantillons négatifs attendus

- Taux de sensibilité

$$r_{SE} = \frac{n_+}{E(n_{+tot})} \times 100\%$$

n_+ : nombre de résultats positifs trouvés

$E(n_{+tot})$: nombre total d'échantillons positifs attendus

- Taux d'exactitude

$$r_{AC} = \frac{n_+ + n_-}{n_{tot}} \times 100\%$$

n_{tot} : nombre total d'échantillons

D'autre part, l'Ifremer réalise des tests de comparaison inter-analystes en histo-cytopathologie³. Il s'agit de tests de reconnaissance d'organismes pathogènes par la lecture de plusieurs lames. La liste des organismes déterminés par les différents opérateurs est comparée à une liste de référence (établie par le responsable technique du laboratoire). Ainsi des pourcentages de concordance sont déterminés pour chaque analyste. Les résultats de ces tests de concordance permettent de calculer des indicateurs de qualité individuel mais également collectif (test statistique de Kappa de Fleiss).

Les données obtenues suite à la lecture des lames (liste d'organismes pathogènes) semblent de même type que la partie qualitative de la donnée hydrobiologique (liste taxonomiques). Il

³ Organisation des tests de comparaison inter-analystes en histo-cytopathologie des mollusques de la Cellule Analytique du Laboratoire de Génétique et de Pathologie de l'Ifremer

pourrait être intéressant de calculer des pourcentages de concordance pour comparer des données qualitatives en hydrobiologie.

5.4. Evaluation des performances

Les normes NF ISO 13528 et NF EN ISO/CEI 17043 précisent qu'il est nécessaire d'établir des critères d'évaluation de la performance. Pour les Z-scores ou scores Zéta par exemple (données quantitatives), les règles d'interprétation de la valeur sont les suivantes selon les deux normes précitées:

- Résultat bon si Z score compris entre -2 et 2 ;
- Résultat suspect (signal d'avertissement) si Z score supérieur à 2 ou inférieur à -2 ;
- Résultat mauvais (signal d'action) si Z score supérieur à 3 ou inférieur à -3.

Stribling *et al.* expliquent également certains critères d'interprétation.

L'importance relative et l'acceptabilité des différentes sources d'erreurs dépendent des objectifs et des données nécessaires : objectifs de qualité de mesure (MQO) et objectifs de qualité de données (DQO). Les objectifs de qualité de mesure pour la fidélité d'un programme de surveillance doivent être fixés à un niveau acceptable pour les utilisateurs de données.

Cette notion est très intéressante, car focalisée sur les objectifs réels de niveau de qualité des données, ce qui n'est pas pris en considération dans les systèmes actuels, où les marges d'écart admissibles sont prédéfinies et fixes, sans aucun lien avec l'usage des données et les besoins.

Ceci pourrait conduire à renforcer les exigences dans certains cas ou certains types d'analyses.

Enfin la norme NF EN 16101 rappelle que l'évaluation des résultats et des modes opératoires statistiques doit être réalisée par des examinateurs ou un jury d'experts indépendants.

5.5. BILAN

Le tableau ci-contre résume les différentes méthodes d'analyse pouvant être utilisées pour les données taxonomiques qualitatives et quantitatives en hydrobiologie, d'après les référentiels applicables recensés.

Analyse / traitement de l'essai d'aptitude	Référence	Type de données	
		Données qualitatives : listes de taxons	Données quantitatives : abondances
Valeur assignée (comparaison avec les résultats)		- valeur de référence déterminée par expert - valeur consensuelle : moyenne robuste des valeurs des participants	- valeur de référence déterminée ou certifiée - valeur consensuelle provenant des participants ou des participants experts
Différence (% ou non)	NF ISO 13528 NF EN ISO/CEI 17043	○	●
Z-score		○	●
Zéta-score		○	●
Ecart normalisés		○	●
Pourcentage de différences dans le dénombrement (PDE)	Stribling <i>et al.</i> 2003	○	●
Pourcentage de désaccords taxonomiques (PTD)		●	○
Pourcentages de taxons communs, en plus ou en moins par rapport à la référence	Archaimbault 2012	●	○
Nombre de résultats + et - (taux de spécificité, de sensibilité et d'exactitude)	XP CEN ISO/TS 22117	●	○
Pourcentages de concordance	Tests histocytologiques	●	○

Figure 5 : Bilan des méthodes d'analyse des résultats d'un essai d'aptitude en hydrobiologie.

On constate que plusieurs méthodes sont utilisables pour l'analyse des données d'un essai d'aptitude mais peu sont normalisées en ce qui concerne les données qualitatives. La majorité des méthodes statistiques décrites dans les référentiels applicables aux essais interlaboratoires en chimie ou en hydrobiologie sont destinées aux données quantitatives, qui ne constituent qu'une partie de la donnée hydrobiologique (abondances). La liste taxonomique est alors généralement transformée en métriques ou indices afin de correspondre à des données quantitatives (richesse taxonomique, par exemple). Nous avons vu, dans l'analyse des spécificités des « mesures hydrobiologiques », que cela peut induire des biais dans la comparaison.

Des méthodes permettant de comparer de façon directe les compositions taxonomiques peuvent exister dans d'autres domaines que la chimie ou l'hydrobiologie (microbiologie, par exemple) mais ne sont pas décrites par des référentiels harmonisés. Il apparaît clairement que les protocoles spécifiques et standardisés restent à développer et à valider, afin de disposer de bases claires et pertinentes pour la comparaison des données hydrobiologiques brutes, utilisant par exemple des analyses multifactorielles.

Il est à noter que des recommandations spécifiques peuvent désormais être incluses dans des référentiels méthodologiques techniques, comme c'est le cas pour l'analyse des résultats d'EIL relatifs à la méthode d'étude des macrophytes en cours d'eau (norme prEN 14184, mars 2012). Dans ce texte, pour les données qualitatives (listes taxonomiques), les préconisations rejoignent les méthodes de calculs de pourcentages de concordance ou de présence/absence citées dans plusieurs documents ci-dessus. Concernant les données quantitatives, il est précisé que le Z-score n'est pas utilisable pour ce type de méthodes hydrobiologiques pour lesquelles l'estimation du recouvrement taxonomique est effectuée sur le terrain et à l'aide de classes de recouvrement (données catégorielles).

6. Exemples d'essais d'aptitude existants

L'offre en essai d'aptitude en hydrobiologie est encore peu importante. Ci-dessous sont présentés les principaux essais qui ont été recensés au niveau national français ou qui sont pratiqués par les laboratoires impliqués dans les programmes de surveillance DCE, ou encore ceux qui sont connus depuis plusieurs années dans les réseaux d'experts européens.

Les essais qui sont organisés, le cas échéant, dans les autres états-membres européens avec une portée principalement nationale n'ont pas été recherchés.

Dans cette approche, la dénomination « d'essais d'aptitude » recouvre une diversité assez large de vocations et de méthode de comparaison, voire d'implication des laboratoires participants.

Le propos est ici d'exposer et de commenter quelques exemples d'essais qui ont déjà été organisés, et de les commenter au regard de l'analyse qui a été développée dans les chapitres précédents. Ces commentaires ne constituent bien évidemment pas un jugement direct sur la conformité, la pertinence ou la qualité intrinsèque de ces essais.

6.1. Essais en cours d'eau

6.1.1. Invertébrés : EIL Aglae

Contenu

Aglae⁴, association d'organisation d'essais interlaboratoires, a mis en place en 2011 un programme d'EIL portant sur les invertébrés benthiques, sous l'intitulé « IBGN et IBG-DCE : liste faunistique et calcul des indices (pas d'étape de prélèvement) ».

Cet essai a d'abord été proposé à titre expérimental en 2010, puis a été validé à partir de la campagne 2011, avec l'accréditation d'AGLAE en tant qu'OCIL⁵ « Hydrobiologie ».

⁴ Voir le site web d'Aglae - <http://www.association-aglae.fr>

C'est en 2013 le seul EIL existant en France sur une méthode hydrobiologique organisé par un OCIL accrédité.

L'essai expérimental de 2010 avait fait l'objet d'une présentation des résultats sous la forme d'une communication au séminaire Aquaref 2010 (Molinier et al., 2011).

L'objectif de l'essai était de « déterminer les compétences des laboratoires en matières de tri, de détermination et de dénombrement des taxons, d'établissement de liste faunistique et de calcul d'indices ». Les laboratoires participants ont reçu 8 à 12 échantillons selon s'ils participaient, respectivement, à l'essai IBGN ou IBG-DCE. Les échantillons sont constitués par un laboratoire spécialisé sur la base de prélèvements sur un site choisi en particulier pour son substrat homogène.

Trois grandes questions ont été posées : « Quelles sont les compétences du laboratoire en termes de dénombrement et de détermination taxonomique ? Quelle est l'aptitude du laboratoire à extraire la variété taxonomique présente dans les échantillons ? Quelle est l'appréciation globale sur les résultats d'analyse des laboratoires ? »

Les résultats obtenus (abondance relative) ont été traités statistiquement taxon par taxon et ont abouti à un z-score généralisé (sur l'ensemble des échantillons analysés). Les différents z-scores généralisés (différents taxons) ont ensuite été cumulés pour donner un indicateur statistique RSz. Des z-scores ont également été calculés à partir des données de variété taxonomique.

Enfin, les laboratoires ont pu être soumis à un classement qualitatif sur la base du taxon retenu pour déterminer le groupe faunistique indicateur (calcul de l'IBGN), sur la variété taxonomique, sur le calcul de l'indice et sur l'attribution d'une classe de qualité.

Commentaires

Cet EIL Invertébrés en cours d'eau concerne uniquement la partie analyse en laboratoire du protocole d'acquisition de la donnée ainsi que le calcul d'un indice biologique. La différenciation des échantillons entre deux méthodes (IBGN et IBGN-DCE) porte sur le nombre d'échantillons à analyser et le type de traitement à appliquer, et non sur la différence de protocole de prélèvement, qui est significative entre les deux méthodes.

Afin de maîtriser la variabilité inter-échantillons, le type de substrat utilisé pour l'échantillonnage vise une homogénéité maximale, en accord avec l'objectif de cet essai. Il ne peut donc pas être représentatif des cas, très répandus en France, des cours d'eau à substrat hétérogène, où la diversité d'habitats, donc de taxons, est potentiellement élevée.

Malgré l'intitulé de l'essai proposé, les méthodes concernées par cet essai ne sont pas en relation directe avec les normes décrivant les méthodes accréditées dans la surveillance DCE. En effet l'« IBG-DCE » ne correspond pas à une méthode normalisée et accréditée, et apporte une confusion méthodologique entre les termes « IBG » et « DCE », puisque l'IBG n'est pas indicateur préconisé dans les méthodes d'évaluation DCE-compatibles. Puisque centré sur le traitement en laboratoire, la mention de la norme XP T90-388 serait sans doute mieux adaptée et précise, si

⁵ OCIL : Organisateur de comparaisons interlaboratoires

c'est bien cette méthode qui est préconisée pour l'essai. De même, pour l'IBGN, la norme AFNOR NF-T90-350 contient à la fois l'échantillonnage, le traitement en laboratoire et le calcul de l'indice. Or, seules les deux dernières étapes sont concernées par l'essai.

Concernant l'analyse des résultats, le calcul de Z-scores a été utilisé pour les données quantitatives (abondances, variété taxonomique). Un classement qualitatif selon le taxon obtenu pour le groupe indicateur est établi.

Ce type d'EIL positionnant les laboratoires les uns par rapport aux autres s'inscrit directement dans le cadre d'essais d'aptitudes requis par les processus d'accréditation. Toutefois, il n'examine pas la totalité des protocoles d'acquisition de données sur lesquels porte l'accréditation ou l'obligation d'agrément ministériel, mais compare les compétences en traitement d'échantillon et détermination taxonomique, deux points essentiels de la démarche.

6.1.2. Diatomées : essai collaboratif des DREAL

Contenu

Un exercice d'intercalibration est organisé par les DREAL depuis 2010 sur l'Indice Biologique Diatomées, à destination des hydrobiologistes des services départementaux et régionaux de l'Etat concernés par l'évaluation de la qualité biologique des cours d'eau. Cet essai collaboratif est basé sur ceux mis en place depuis 2002 par les DIREN, il est ouvert à d'autres laboratoires depuis la campagne 2013. Regroupant tous les laboratoires pendant une semaine sur un site différent chaque année, il est encadré par un expert diatomiste (Luc Ector, CRP Gabriel Lippmann- Luxembourg).

Cet exercice permet la comparaison des résultats obtenus par les différents participants au niveau :

- de l'échantillonnage,
- de la préparation des échantillons,
- de la détermination.

Chaque participant réalise un prélèvement sur quatre mêmes stations. Le prélèvement permet de préparer un échantillon individuel et un « pot commun » pour analyse d'un échantillon commun. Chaque participant présente donc 2 résultats par station.

Les listes taxonomiques ainsi que les notes IBD obtenues sont comparées par des méthodes simples (répartition des résultats représentés sur des box-plots), par rapport aux résultats de référence obtenus par l'expert.

Un de ses principaux objectifs est d'homogénéiser et d'améliorer les compétences des laboratoires participant en insistant sur les aspects « échange et formation », mais aussi de tester et améliorer les référentiels de détermination taxonomique (guide méthodologique, atlas iconographiques).

Commentaires

Ce type d'essai collaboratif n'est pas réellement un essai d'aptitude puisque sa principale visée est la formation, le perfectionnement, l'échange entre diatomistes. Sa vocation ne vise pas un

classement de la capacité de chaque laboratoire, mais des objectifs multiples. A ce titre, il répond à l'objectif d'amélioration de la qualité des données produites, mais ses résultats ne sont pas directement utilisables pour situer les performances d'un laboratoire par rapport aux autres. Il ne permet notamment pas de calculer de Z-score ou d'autres métriques standard de comparaisons des compétences des opérateurs.

Cet essai collaboratif concerne le protocole d'acquisition de donnée dans son ensemble, il suit la réalisation du protocole « IBD » tel qu'il est actuellement normalisé (NF T90-354) et soumis à l'accréditation.

6.1.3. Diatomées : comparaison interlaboratoire AGLAE

En tant qu'OCIL, AGLAE organise en 2013 un essai interlaboratoire portant sur le protocole de la norme de l'Indice Biologique Diatomées, pour une première expérimentation. Cet essai est basé sur un échantillon prélevé par un laboratoire accrédité pour ce protocole, et envoyé à tous les laboratoires participants.

Cet essai porte donc uniquement sur les phases de traitement des échantillons en laboratoire : préparation des échantillons, détermination taxonomique et comptage. Les phases d'exploitation des résultats pour le calcul de l'IBD sont également considérées, puisque l'essai intègre la saisie des résultats sous un logiciel dédié (OMNIDIA) et le calcul de l'indice.

L'objectif s'inscrit, comme pour l'essai « invertébrés », dans le cadre standardisé de la comparaison des performances des laboratoires sur les phases réalisées en laboratoire.

A la date de rédaction de la présente synthèse, le détail de l'exploitation de cet essai n'est pas connu.

6.2. Essais en plans d'eau

6.2.1. Phytoplancton

Contenu

Il n'existe actuellement pas d'EIL au niveau français pour cet élément biologique. En revanche, deux EIL ont été mis en place au niveau européen par des organismes de recherche et développement accrédités en tant qu'OCIL :

- ATT en Allemagne - <http://www.planktonforum.eu>
- SYKE en Finlande - <http://www.syke.fi>

L'EIL proposé par ATT comprend 3 composantes :

- Les participants doivent estimer la concentration, la concentration volumique, et le diamètre des microparticules des 3 échantillons synthétiques sur des cellules de comptage de référence.
- Ils doivent également estimer les abondances et les biovolumes de 3 échantillons naturels (abondance totale et pour chaque classe, biovolume total et pour chaque classe).
 - ⇒ Pour toutes ces données quantitatives, des comparaisons statistiques sont effectuées : calcul d'écart-types et de Z-scores.

- ⇒ Un lien avec la stratégie de comptage employée par les participants est également évalué par le calcul de coefficients de variation.
- Enfin, des vidéos sont présentées aux participants afin de tester les compétences en détermination taxonomique.
 - ⇒ Un score est calculé pour chaque laboratoire (non comparé statistiquement) pour toutes les vidéos (20 vidéos) et ces 20 scores sont additionnés.

L'EIL organisé par SYKE comprend trois phases de l'analyse du phytoplancton :

- Test d'identification des espèces sur 20 vidéos. Les échantillons filmés sont des échantillons préparés, conservés dans du Lugol.
 - ⇒ L'analyse des résultats du test d'identification des taxons est réalisée à l'aide d'un score allant de 0 à 3 (principe identique à celui de l'EIL d'ATT).
- Test de comptage sur 25 vidéos représentant 25 champs dans un microscope. Les échantillons filmés sont des échantillons composites (préparés en mélangeant du phytoplancton naturel de lac et du phytoplancton de culture). Il est recommandé aux participants d'utiliser la méthode de comptage décrite dans la norme EN 15204 (2006).
- Test de mesure du diamètre des cellules. Des échantillons composites sont distribués aux participants.
 - ⇒ Des analyses statistiques sont réalisées selon la norme ISO 13528 (2005) et notamment par l'utilisation du z-score pour les tests de comptage et de mesure.

On peut noter que SYKE propose également un test de capacité concernant l'identification d'invertébrés benthiques, sur le même principe.

Commentaires

Ces deux EIL concernent uniquement la partie analyse en laboratoire du protocole d'acquisition de la donnée. La détermination est réalisée sur échantillons synthétiques et naturels ainsi que sur des vidéos. Concernant le phytoplancton, on peut considérer que les phases réalisées en laboratoire sont celles qui ont le plus de poids dans le résultat final, si on considère l'ensemble de la méthode. En effet, la phase de terrain consiste en un prélèvement d'eau « simple », qui doit être encadré en tant que tel pour garantir un échantillonnage conforme aux objectifs, au regard des caractéristiques particulières de la colonne d'eau d'un lac.

Les données quantitatives de ces deux EIL sont analysées statistiquement à l'aide du calcul de Z-scores, les données qualitatives étant décrites et non analysées statistiquement.

6.3. Essais en eaux littorales

Les essais portant sur les méthodes « eaux littorales » sont encore extrêmement peu nombreux. Seules deux opérations de ce type ont été menées par l'Ifremer, à titre d'expérimentation. Ces actions ayant apporté des résultats très positifs, elles devraient se poursuivre, sur d'autres méthodes et avec d'autres opérateurs.

6.3.1. Macro-algues

Contenu

Dans le cadre des actions d'Aquaref, Ifremer a organisé en 2012, en collaboration avec le LEBHAM - IUEM - UB,O Laboratoire d'Ecophysiologie et de Biotechnologie des Halophytes et des Algues de l'Université de Bretagne Occidentale à Brest, un premier essai collaboratif concernant le relevé et le calcul d'indicateurs sur les macro-algues intertidales.

Un des objectifs premiers de cet essai collaboratif est la définition d'une méthodologie d'EIL qui pourrait servir de base à l'organisation de tests par des OCIL. Le deuxième objectif, opérationnel, est la vérification de la bonne application des protocoles par les différents participants afin que les résultats soient comparables sur l'ensemble du littoral surveillé et ainsi permettent l'optimisation et la standardisation des méthodes utilisées. Cette opération a donc une vocation multiple, en mettant en avant la formation et l'échange entre les opérateurs et pas uniquement la comparaison des résultats de chaque laboratoire.

Cet EIL comprend une comparaison des résultats obtenus au niveau :

- De la phase terrain : une partie de l'échantillonnage (le relevé)
 - estimation des surfaces végétalisables et du pourcentage de couvert végétal (par ceinture à l'intérieur des limites du site) ;
 - reconnaissance des taxons (identification des espèces caractéristiques et opportunistes et estimation de leur pourcentage de recouvrement par quadrat) ;
 - calcul de métriques.Les ceintures et quadrats ont été disposés par l'équipe de référence et tous les participants ont travaillé sur les mêmes zones.
- De la phase laboratoire :
 - calcul de l'indice ;
 - classement de la masse d'eau.

L'analyse des résultats de cet EIL est présentée dans le rapport intitulé « Essais interlaboratoires sur les macro-algues intertidales en milieu marin » (Miossec 2013).

En 2013, Ifremer a organisé un autre essai de ce type, portant sur les algues subtidales, regroupant les laboratoires impliqués dans le suivi de cet élément biologique. Les résultats de cette action, partie du programme Aquaref 2013, seront diffusés dans le rapport technique spécifique en 2014.

Commentaires

Comme pour l'essai diatomées en cours d'eau, cet essai est collaboratif, regroupant l'ensemble des équipes impliquées dans le développement et dans l'application de la méthode « Algues intertidales ». Il vise également l'évaluation des compétences des participants dans ses objectifs principaux.

Cet essai concerne une grande partie du protocole d'acquisition de la donnée (phase terrain et laboratoire), seule la partie de mise en place des filets pour la délimitation des quadrats n'est pas évaluée. De plus, les opérateurs ont bénéficié de l'aide d'un expert (objectif de formation) lors des relevés.

Ce type d'essai collaboratif est exploratoire, puisque la méthode elle-même est encore en cours de validation. Il ne permet pas de comparer les performances des laboratoires car il mêle deux objectifs de validation de la méthode et de comparaison des pratiques opérateurs. Il s'agit donc principalement de regrouper les intervenants de ces mesures pour homogénéiser les niveaux de pratique, répondre aux interrogations des opérateurs, et en tirer des enseignements quant aux besoins de formation, de documents d'encadrement (guide méthodologiques), et de principes d'organisation d'EIL plus formels.

6.3.2. Phytoplancton

Contenu

Ifremer a mis en place en 2007 un EIL sur le phytoplancton (comptage et détermination) dans le cadre du réseau REPHY. Les résultats ont été diffusés en 2009 (Grossel 2009).

L'objectif de cet EIL était de « permettre aux laboratoires participants de se situer vis à vis de l'ensemble des résultats obtenus et de mener le cas échéant la réflexion en vue de l'optimisation de leurs pratiques et l'approfondissement de leur expertise taxinomique ». Cet essai peut donc également être considéré comme essai collaboratif puisqu'il a pour objectif la formation des participants.

Des échantillons issus de populations naturelles de phytoplancton ont été fournis aux laboratoires participants. Les analystes ont observé ces échantillons suivant la méthode d'Utermöhl.

L'analyse statistique a permis la détermination d'un z-score pour les résultats quantitatifs : nombre de cellules comptées, comptages par taxons (pour les espèces particulières). Une analyse descriptive a également été faite sur les comptages spécifiques et les questions taxonomiques.

Commentaires

Dans ce cas également, cet essai est collaboratif de part son objectif de formation. Il intègre toutefois l'évaluation des compétences des participants dans ses objectifs principaux, en calculant un indicateur de performance comparée des laboratoires.

Les données quantitatives et qualitatives ont été analysées, le comptage à l'aide de Z-scores, les listes taxonomiques de manière descriptive.

Cet essai concerne uniquement la partie « analyse » du protocole d'acquisition de la donnée. Comme nous l'avons noté pour le phytoplancton lacustre, on peut toutefois considérer que

l'examen des phases réalisées en laboratoire répond aux objectifs d'évaluation des compétences hydrobiologiques pour cet élément « phytoplancton ».

7. Synthèse, perspectives

La synthèse de la réflexion menée sur l'adaptation de la notion d'essais interlaboratoires aux méthodes hydrobiologiques fait apparaître plusieurs points principaux, liés à la nature des « mesures » hydrobiologiques, à celle des résultats issus de telles mesures, et aux objectifs de ces essais :

- **Les notions d'échantillonnage - prélèvement et de traitement en laboratoire ne sont pas dissociables *a priori***, dans la majorité des méthodes utilisées en évaluation hydrobiologique. En effet, l'approche peut comprendre un plan d'échantillonnage complexe étroitement lié à la station ou un relevé biologique réalisé directement sur site pour sa plus grande partie. La distinction fonctionnelle des deux phases de terrain et de laboratoire n'est donc généralement pas pertinente au regard de la cohérence d'ensemble du protocole. Quelques cas peuvent toutefois se prêter à un essai focalisé uniquement sur le travail en laboratoire (préparation d'échantillons, déterminations, comptages). C'est le cas en particulier des méthodes « phytoplancton » et « diatomées », pour lesquelles la phase de terrain est un prélèvement qui peut être considéré soit comme relevant d'autres types d'essais à mettre en place le cas échéant (prélèvement d'eau pour le phytoplancton), soit comme influençant relativement peu le résultat final (diatomées). Il n'est donc pas possible de définir le contenu type d'un essai comparatif interlaboratoires pour l'ensemble des méthodes hydrobiologiques. Cette indissociabilité des phases techniques des méthodes hydrobiologiques devrait être précisée par l'analyse des sources d'incertitude de chaque protocole. En effet, pour des raisons pragmatiques de rapport coût/information, il peut être envisageable que des essais interlaboratoires ne visent que certaines phases, à la condition de couvrir celles qui apportent la plus forte incertitude dans le résultat. Cet ajustement ne pourra être complètement validé que lorsqu'on disposera d'une connaissance suffisante sur ces sources d'incertitude et leur quantification.

Dans la majorité des cas, les essais visant uniquement la partie analyse ne permettent donc pas de tester la capacité des opérateurs à appliquer une méthode complète. Toutefois, ces types d'essais peuvent avoir un intérêt pour les laboratoires pour tester certaines capacités, en particulier la détermination taxonomique, cœur des méthodes hydrobiologiques.

- **Concernant les procédures formelles d'accréditation**, il faut rappeler que l'accréditation, telle qu'elle est organisée actuellement, est délivrée sur la base d'un référentiel technique (en complément de l'aspect « qualité »), en général une norme. Au regard du besoin de positionner les performances des opérateurs dans le cadre de cette démarche (préconisation dans la norme ISO 17025, obligation dans l'arrêté Agrément des laboratoires), il semble indispensable que les essais interlaboratoires intègrent l'ensemble de la méthode en question. On peut alors considérer que les essais qui n'examinent qu'une partie d'un protocole normalisé, ou qu'une seule norme lorsque plusieurs sont données comme indissociables pour la

mise en œuvre d'une méthode, ne permettent pas une évaluation pertinente des laboratoires par rapport à l'objectif global.

Actuellement, les essais existants organisés par des OCIL principalement en vue de ces démarches d'accréditation ne considèrent que les phases de traitement en laboratoire, voire uniquement la reconnaissance d'espèces. Ils ne répondent donc pas entièrement aux besoins. Il est cependant nécessaire que l'accréditation puisse s'appuyer sur des procédures objectives et généralisées pour assurer les compétences des structures accréditées. Les exercices existants qui représentent des pistes exploratoires prometteuses doivent être consolidées pour répondre pleinement aux besoins actuels.

- **La nature des méthodes hydrobiologiques engendre des difficultés intrinsèques à ces protocoles et à leur objet d'étude.** L'échantillonnage est dans certains cas semi-destructif, et rend donc impossible les essais de répétabilité ou de comparaison de plusieurs opérateurs dans un court laps de temps sur un même site. La variabilité des habitats aquatiques introduit de même des biais plus ou moins contrôlables dans un échantillonnage comparatif. De tels essais ne peuvent donc être envisagés que pour les protocoles qui ne sont pas significativement impactés par ces éléments.

Un autre aspect découlant des caractéristiques écologiques des milieux inventoriés est l'écorégionalisation des habitats et des taxons qui servent de base à l'étude. Dans le cadre des procédures qualité liées à la surveillance DCE, la portée des EIL doit être au minimum nationale. La comparaison des performances peut alors être faussée par le fait que les opérateurs ont rarement un niveau de compétence identique pour toutes les écorégions du territoire national, ce qui est compréhensible vu les particularités des habitats aquatiques et des espèces potentiellement rencontrés à cette échelle territoriale très variée. Pour être significatif, l'EIL devrait permettre d'évaluer la performance des opérateurs dans des conditions comparables à celles dans lesquelles ils opèrent en routine. Cela peut conduire à limiter les tests à des habitats ou taxons communs à l'ensemble des domaines biogéographiques nationaux, ce qui risque d'induire d'autres inconvénients (pas de tests sur des taxons particuliers très bioindicateurs). Remarquons que, dans la pratique, de nombreux opérateurs parmi les laboratoires privés interviennent indifféremment sur la totalité du territoire national, selon l'attribution des lots des marchés publics.

Cet aspect de la répartition des taxons rend également difficile à envisager l'application de tels essais dans les **départements d'Outre-Mer**, même si une partie des taxons est assez cosmopolite chez certains groupes (diatomées) ou comparable à un niveau taxonomique moins précis (les familles d'invertébrés). Pour ces territoires tropicaux, la question devra être examinée au cas par cas, mais il ne sera *a priori* pas possible d'organiser des EIL sur les mêmes bases. D'autant plus que le nombre de laboratoires opérant dans ces départements est extrêmement réduit (souvent un seul laboratoire par DOM), ce qui rend toute démarche comparative sans objet.

- **La performance des laboratoires doit être évaluée sur la production des résultats bruts.** Ce résultat correspond en effet à ce qui est produit en suivant une norme technique (pour celles qui ont d'ores et déjà été mises en compatibilité avec les préconisations de la DCE), et fait l'objet d'une accréditation. Le calcul de métadonnées (métriques, indices, classes d'état)

n'est pas du même domaine technique ni du ressort des mêmes intervenants dans le schéma d'évaluation DCE. De plus, ces métadonnées peuvent masquer certains défauts des données de base, et ne pas réagir à des erreurs ou à l'incomplétude des données. Ces données ayant aussi vocation à être utilisées dans d'autres cadres ou avec d'autres règles d'exploitation (métriques d'évaluation et de diagnostic en cours de développement), il est indispensable que ce soit ces résultats eux-mêmes qui soient l'objet des contrôles de qualité.

Cette contrainte implique que l'exploitation des résultats des EIL soit adaptée pour montrer les écarts dans ces données brutes. Rappelons que le résultat d'une « mesure hydrobiologique » au sens de la DCE est dans la majorité des cas une liste faunistique ou floristique assortie des abondances taxonomiques. Les critères caractérisant ce résultat concernent donc la complétude de la liste (l'opérateur a-t-il vu tous les taxons présents ?), sa justesse (l'identification de chaque taxon est-elle correcte ?), et l'estimation des abondances (le comptage de chaque taxon, ou l'estimation de son recouvrement, est-il juste ?).

- Les comparaisons interlaboratoires peuvent être envisagées sous plusieurs aspects.

Jusqu'à présent, la notion recouvrait surtout les EIL de performance, car c'est le besoin qui est formulé depuis peu de temps de façon réglementaire, en particulier par l'arrêté Agrément des laboratoires d'octobre 2011. L'analyse ci-dessus montre les difficultés que la construction de ces tests peut soulever.

D'autres essais se mettent en place ou se développent, sous la forme d'essais collaboratifs. L'intérêt est évident, car ils répondent à plusieurs besoins, tout en contribuant directement à l'amélioration des conditions dans lesquelles sont acquises les données de surveillance, ce qui est l'objectif final de ces démarches « qualité ». Ces opérations intègrent l'échange entre les opérateurs sur les pratiques, la formation, l'encadrement de la mise en œuvre des protocoles par les personnes référentes. Ce type d'essai induit donc directement à la fois l'augmentation effective des compétences des opérateurs (qualité des données elles-mêmes) et l'homogénéisation des pratiques (qualité des bases de données nationales), comme l'illustre le schéma de la figure 6.

Ces essais ne sont pas encore bien adaptés à la comparaison des laboratoires entre eux. Il serait intéressant de rechercher des pistes d'évolution dans ce sens, afin que ces comparaisons puissent être directement valorisées dans le cadre de démarches d'accréditation, par exemple.

Même s'ils peuvent paraître plus lourds à mettre en place, tant pour les organisateurs que pour les participants, ils sont probablement le moyen le plus efficace, dans le contexte actuel, pour atteindre des objectifs d'amélioration de la qualité des données hydrobiologiques, au meilleur rapport coût/bénéfice. Ils peuvent aussi s'inscrire à terme dans les parcours de formation qui

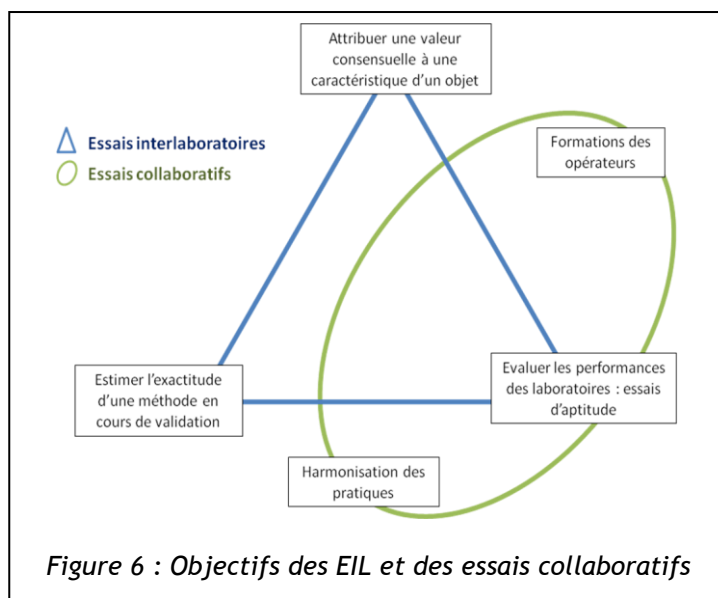


Figure 6 : Objectifs des EIL et des essais collaboratifs

seront développés sur les méthodes DCE, ce qui serait un moyen de mettre en cohérence plusieurs aspects de l'encadrement des opérateurs, qui visent un même objectif. Ces essais collaboratifs seraient à privilégier et à développer, car ils répondent complètement à l'objectif des comparaisons inter-opérateurs, et apportent des éléments de réponse nécessaires pour poursuivre et améliorer la démarche globale.

- La demande en essais d'aptitude est devenue forte du fait de la nécessité formelle pour les laboratoires de réaliser ces essais dans le cadre de l'accréditation et de l'agrément, mais **cette notion est encore très jeune en hydrobiologie**. Si de tels essais doivent être développés, il sera nécessaire de mieux examiner les liens avec les essais d'aptitude menés en chimie afin de travailler sur la construction conjointe de ces essais, le cas échéant. Les tests organisés pour les analyses chimiques peuvent d'ailleurs être remis en question, sur certains aspects, au vu de la réflexion engagée pour l'hydrobiologie (importance des phases d'échantillonnage, par exemple, dans l'évaluation de l'état des masses d'eau elles-mêmes). Dans la même optique, il pourrait être utile d'engager une collaboration avec les OCIL pour développer ces tests. Toutefois, les OCIL n'ont encore qu'une expérience très réduite de ce domaine de la comparaison interlaboratoires en hydrobiologie, et sont très peu nombreux à s'y investir (un seul en France actuellement). De plus, le principe des essais menés par ces organismes, dont les résultats ne sont pas diffusés en dehors du cercle des laboratoires participants, est un obstacle à l'utilisation des enseignements qui en sont issus.

La réflexion pour la mise en place de procédures de validation des compétences des laboratoires d'hydrobiologie doit se concevoir de façon globale. Il s'agit en effet de définir les moyens les mieux adaptés pour atteindre l'objectif général de maîtrise et d'amélioration des données hydrobiologiques. La formation est probablement un point essentiel de ce système, puisqu'il est reconnu que la compétence et l'expérience de l'opérateur hydrobiologiste lui-même sont des éléments-clé incontournables de la qualité des résultats, plus que dans d'autres domaines de la surveillance (chimie). Cela est d'autant plus vrai actuellement que la plupart des méthodes mises en œuvre dans les réseaux DCE sont récentes, voire encore expérimentales pour certaines ; les opérateurs ne sont donc pas toujours familiarisés avec leur application. La conception des comparaisons interlaboratoires devra prendre en considération cette dimension, et sans doute privilégier les essais collaboratifs aux EIL plus formels sanctionnant uniquement une performance.

Au vu de l'expertise hydrobiologique des concepteurs des méthodes d'évaluation, et même si cela paraît une évidence, **il est essentiel que les exigences normatives ou réglementaires garantissent l'amélioration réelle de la qualité et de la robustesse des données recueillies *in fine* en routine**, et ne visent pas uniquement un simple positionnement formel des laboratoires entre eux selon des critères de compétence sur certains points particuliers ou des normes « qualité. En effet, ces référentiels ne mettent pas toujours suffisamment en avant les aspects « métier » et l'expérience, voire l'expertise, des opérateurs dans la chaîne complète d'acquisition de données. Ceci est d'autant plus à considérer que les laboratoires d'hydrobiologie, à la différence des laboratoires d'analyse chimique, sont pour partie des structures de petite voire très petite taille, pour lesquelles la démarche qualité peut être une

charge très lourde. Il est alors primordial que les prescriptions obligatoires soient basées sur des procédures à rapport effort/bénéfice collectif efficient.

En tout état de cause, même si cela peut nécessiter de revisiter certains aspects des exigences réglementaires ou normatives actuelles, l'objectif des essais interlaboratoires devra bien suivre la philosophie qui a généralement prévalu aux dispositions énoncées dans ces textes : améliorer et fiabiliser la qualité, la représentativité et l'homogénéité finale des données acquises dans les réseaux de surveillance, dans l'optique générale qui est l'évaluation de l'état écologique des masses d'eau.

Travaux et ouvrages consultés

- Afnor (1994). NF ISO 5725-1 (1994-12-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 1 : principes généraux et définitions. .
- Afnor (1994). NF ISO 5725-2 (1994-12-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 2 : méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée. .
- Afnor (1994). NF ISO 5725-3 (1994-12-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 3 : mesures intermédiaires de la fidélité d'une méthode de mesure normalisée.
- Afnor (1994). NF ISO 5725-4 (1994-12-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 4 : méthodes de base pour la détermination de la justesse d'une méthode de mesure normalisée.
- Afnor (1994). NF ISO 5725-6 (1994-12-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 6 : utilisation dans la pratique des valeurs d'exactitude.
- Afnor (1998). NF ISO 5725-5 (1998-10-01). Application de la statistique - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 5 : méthodes alternatives pour la détermination de la fidélité d'une méthode de mesure normalisée.
- Afnor (2004). NF T90-350 (2004-03-01). Qualité de l'eau - Détermination de l'indice biologique global normalisé (IBGN).
- Afnor (2005). NF EN ISO/CEI 17025 (2005-09-01). Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais.
- Afnor (2005). NF ISO 13528 (2005-12-01). Méthodes statistiques utilisées dans les essais d'aptitude par comparaisons interlaboratoires.
- Afnor (2006). ISO/CEI GUIDE 35:2006 (2006-01-01). Matériaux de référence - Principes généraux et statistiques pour la certification.
- Afnor (2010). NF EN ISO-CEI 17043 (2010-04-01). Évaluation de la conformité - Exigences générales concernant les essais d'aptitude.
- Afnor (2010). XP T90-388 (2010-06-01). Qualité de l'eau - Traitement au laboratoire d'échantillons contenant des macro-invertébrés de cours d'eau.
- Afnor (2012). NF EN 16101 (2012-12-01). Qualité de l'eau - Guide pour les études comparatives interlaboratoires ayant pour objet l'évaluation écologique. .
- Afnor (2012). prEN 14184 (en cours d'approbation, parution prochaine). Qualité de l'eau - Guide pour l'étude des macrophytes aquatiques dans les cours d'eaux.
- Aquaref (2011). Séminaire - Surveillance des milieux aquatiques - De la maîtrise de la qualité des données à la prise de décision : quid des incertitudes ? Paris, décembre 2011. Communications téléchargeables sur le site web d'Aquaref.
- Aquaref (2013). Séminaire - Amélioration de la qualité des données de surveillance en chimie et en hydrobiologie. Paris, 19 Juin 2013. Communications téléchargeables sur le site web d'Aquaref.
- Archambault, V. (2012). Evaluation de l'incertitude liée à l'application du protocole d'échantillonnage des macro-invertébrés en cours d'eau peu profonds selon la norme Norme XP T90-333. Rapport d'étape Onema année 2011.
- Dagens, N., Chauvin, C. (2013). L'incertitude en hydrobiologie - Rapport Aquaref.
- Ghestem, J., Champion, R., Fisicaro, P. (2009). Rapport de l'essai collaboratif sur l'échantillonnage en eau souterraine - BRGM, LNE - Rapport Aquaref.
- Grossel, H. (2009). "Essai interlaboratoires pour le dénombrement et la taxinomie du phytoplancton marin-Convention 2009-Action 1."

- JORF (2012). Avis du 04 février 2012 relatif aux méthodes des couples « élément de qualité biologique - méthode » sur lesquels porte l'agrément des laboratoires. JORF n°0030 du 4 février 2012 page 2085 texte n° 98. .
- MEDDE (2007). Circulaire DCE 2007/22 du 11 avril 2007 relative au protocole de prélèvement et de traitement des échantillons des invertébrés pour la mise en oeuvre du programme de surveillance sur cours d'eau. Bulletin Officiel du Ministère de l'écologie, du développement et de l'aménagement durables n° 2007/12 du 30 juin 2007. .
- MEDDTL (2011). Arrêté du 27 octobre 2011 portant modalités d'agrément des laboratoires effectuant des analyses dans le domaine de l'eau et des milieux aquatiques au titre du code de l'environnement. JORF n°0260 du 9 novembre 2011 page 18819 texte n°6.
- MEEDDM (2010). Arrêté du 25 janvier 2010 établissant le programme de surveillance de l'état des eaux en application de l'article R. 212-22 du code de l'environnement. JORF n°0046 du 24 février 2010 page 3406 texte n° 8. .
- MEEDDM (2010). Arrêté du 26 juillet 2010 approuvant le schéma national des données sur l'eau. JORF n°0195 du 24 août 2010 page 15272 texte n° 2. .
- Miossec, L., Soudant, D., Ar Gall, E., Le Duff, M. (2013). "Essai interlaboratoire sur les Macroalgues intertidales en milieu marin - Une première approche de protocole."
- Molinier, O., Guarini, P., Fontan, B., Lambry, M., Blanchard, M., Bourgeois, S. (2011). "Essai interlaboratoires en hydrobiologie. Restitution d'une expérience concrète réalisée en 2010". Séminaire AQUAREF "De la maîtrise de la qualité des données à la prise de décision : quid des incertitudes ?" 7 décembre 2011, Paris.
- Prygiel, J., Carpentier, P., Almeida, S., Coste, M., Druart, J. C., Ector, L., Guillard, D., Honoré, M. A., Iserentant, R., Ledeganck, P., Lalanne-Cassou, C., Lesniak, C., Mercier, I., Moncaut, P., Nazart, M., Nouchet, N., Peres, F., Peeters, V., Rimet, F., Rumeau, A., Sabater, S., Straub, F., Torrisi, M., Tudesque, L., Van de Vijver, B., Vidal, H., Vizinet, J., Zydek, N. (2002). "Determination of the biological diatom index (IBD NF T 90-354): Results of an intercomparison exercise." *Journal of Applied Phycology* **14**(1): 27-39.
- Schilling, P., Powilleit, M., Uhlig, S. (2006). "Macrozoobenthos interlaboratory comparison on taxonomical identification and counting of marine invertebrates in artificial sediment samples including testing various statistical methods of data evaluation." *Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement* **11**(8-9): 422-429.
- Stribling, J. B., Moulton, S. R., Lester, G. T. (2003). "Determining the quality of taxonomic data." *Journal of the North American Benthological Society* **22**(4): 621-631.
- Vuorio, K., Lepistö, L., Holopainen, A. L. (2007). "Intercalibrations of freshwater phytoplankton analyses." *Boreal Environment Research* **12**(5): 561-569.