



HAL
open science

Evaluation diatomique des Très Grands Cours d'Eau (TGCE) de France et intercalibration européenne : Rapport Final

D. Carayon, Juliette Tison-Rosebery, Michel Coste, François Delmas

► **To cite this version:**

D. Carayon, Juliette Tison-Rosebery, Michel Coste, François Delmas. Evaluation diatomique des Très Grands Cours d'Eau (TGCE) de France et intercalibration européenne : Rapport Final. [Rapport de recherche] irstea. 2018, pp.131. hal-02608976

HAL Id: hal-02608976

<https://hal.inrae.fr/hal-02608976>

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.



Programme 2016/2018

Fiche-Action 23 : Bio-indication Végétale en Cours d'Eau (Macrophytes et Phytobenthos)



Elaboration d'un dispositif d'évaluation diatomique des Très Grands Cours d'Eau (TGCE) de France et Intercalibration Européenne

Rapport final

Co-Auteurs :

David Carayon & Julie Guéguen, Juliette Rosebery, Michel Coste, François Delmas

Irstea, Centre de Bordeaux, Unité de Recherche EABX /
Equipe ECOVEA

50, Avenue de Verdun, Gazinet, 33 612 CESTAS Cedex

V₁ (Décembre 2018)

- **AUTEURS**

David CARAYON, Ingénieur d'Etude (Irstea Bordeaux) E-mail : david.carayon@irstea.fr

Julie GUEGUEN, Ingénieure d'Etude (ex-Irstea Bordeaux) E-mail : julie.gueguen@inra.fr

Juliette ROSEBERY, Chargée de Recherches (Irstea Bordeaux) E-mail : juliette.rosebery@irstea.fr

Michel COSTE, Directeur de Recherches émérite (Irstea Bordeaux) E-mail : michel.coste@irstea.fr

François DELMAS, Ingénieur-Chercheur IPEF (Irstea Bordeaux) E-mail : francois.delmas@irstea.fr

- **CORRESPONDANTS**

AFB : Yorick REYJOL. E-mail : yorick.reyjol@afbiodiversite.fr

AFB : Bénédicte AUGÉARD. E-mail : benedicte.augeard@afbiodiversite.fr

- **AUTRES CONTRIBUTEURS**

Néant

Droits d'usage : accès libre

Niveau géographique : national

Couverture géographique : France (Très grands Cours d'Eau des 5 grands bassins hydrographiques)

Niveau de lecture : experts, professionnels et gestionnaires dans le domaine de l'évaluation écologique des hydrosystèmes (cours d'eau)



Elaboration d'un dispositif d'évaluation diatomique
des TGCE de France
& Intercalibration Européenne

Rapport final

David Carayon & Julie Guéguen, Juliette Rosebery,
Michel Coste, François Delmas

AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ

Établissement public du ministère de l'Environnement

- **RESUME**

La mise en œuvre de la Directive Cadre sur l'Eau requiert l'évaluation régulière de l'Etat Ecologique des masses d'eau, dans le cadre de plans de gestion de 6 ans. L'objectif poursuivi est d'éviter la dégradation de l'Etat Ecologique, ou de restaurer le Bon Etat Ecologique à l'échéance prévue dans les SDAGE, et au plus tard à fin 2027. Dans cet objectif, une première Circulaire Nationale Evaluation a été publiée en Juillet 2005. Ce texte a depuis fait l'objet de plusieurs réactualisations dont la précédente (Arrêté Evaluation du 27 Juillet 2015) avait force d'application sur la plupart de nos types de cours d'eau nationaux. Cependant, les très grands cours d'eau au sens de la DCE (surface cumulée de bassin versant dépassant les 10 000 km² à la station d'observation) n'y étaient pas encore couverts par un dispositif d'évaluation cohérent du fait de certaines difficultés spécifiques rencontrées à notre échelle nationale comme à l'échelle Européenne, dont l'absence complète de références vraies sur de tels systèmes relativement à fortement anthropisés.

La présente action, soutenue dans le cadre de la fiche-action Irstea-ONEMA N° 23 « Bio-indication végétale en cours d'eau » (2016-2018), avait pour but de combler au plus vite ce déficit réglementaire en proposant un nouveau dispositif national *ad hoc*, si possible harmonisé d'emblée avec les autres Etats-Membres Européens. Dans cette optique, le travail a d'une part consisté à rassembler et à analyser les données comparables obtenues selon un mode suffisamment homogène sur ces types d'hydrosystèmes dans le cadre des réseaux de surveillance nationaux (données IBD₂₀₀₇ recueillies de 2007 à 2013) ; il a d'autre part couvert la participation à un exercice européen d'intercalibration des systèmes d'évaluation de tels cours d'eau, dont l'organisation était concomitante à ce chantier national.

Un premier exercice d'intercalibration européenne avait été réalisé en 2012-2013 par le X-GIG "Large Rivers" sur divers maillons-clés biologiques. Dans ce contexte un peu prématuré par rapport aux avancées de notre propre réflexion nationale, la France a participé par la fourniture de données et en tant qu'observateur externe, mais ne s'est pas prêtée à l'intercalibration. Durant ce premier exercice, les pays participants ont pu se mettre d'accord sur un niveau de référence par défaut, qui a été défini sur un mode consensuel pour ce type d'hydrosystèmes, et ont ensuite intercalibré leurs méthodes nationales sur cette base. Une relation pression-impact entre la teneur en PO₄ des eaux et les réponses de la métrique commune d'inter-calibration (ICM diatomique) a été établie, à partir de laquelle une vision partagée du bon état des très grands cours d'eau a été définie.

Entre 2015 et 2017, la France a donc pris part au deuxième exercice d'intercalibration intitulé "Fit-in exercise", mené sous la coordination conjointe de la Bulgarie et de l'Autriche avec la participation de 8 États membres différents, et qui s'est appuyé sur les résultats précédemment obtenus par le "LR" X-GIG. Les résultats de cet exercice collectif, communiqués par oral et de manière informelle courant 2017, ont été officialisés par la signature de la décision UE 2018/229 du 12-02-2018, validant du même coup le prototype de système d'évaluation proposé par la France dans le cadre de cet exercice. L'intégration du nouveau système dans un modificatif de l'Arrêté "Evaluation" français, proposée entre octobre 2017 et février 2018, est devenue officielle suite à la signature de ce nouveau texte en date du 27 juillet 2018 (Arrêté paru au JORF du 28 Août 2018). Ce nouveau système d'évaluation se base sur 3 grilles différentes qui permettent d'évaluer de manière différenciée et judicieuse 3 types de grands fleuves (plaines peu minéralisées, plaines minéralisées; grands cours d'eau directement issus de massifs montagneux).

- **MOTS CLES (THEMATIQUE ET GEOGRAPHIQUE)**

DCE ; DIRECTIVE COMMUNAUTAIRE SUR L'EAU ; DIATOMEES BENTHIQUES ; IBD₂₀₀₇ ; SYSTEME D'EVALUATION; TRES GRANDS COURS D'EAU; FRANCE; ETAT ECOLOGIQUE



Elaboration d'un dispositif d'évaluation diatomique
des TGCE de France
& Intercalibration Européenne

Rapport final

David Carayon & Julie Guéguen, Juliette Rosebery,
Michel Coste, François Delmas

AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ

Établissement public du ministère de l'Environnement

- **TITLE**

Development of a diatom-based assessment system for French large rivers and European intercalibration

- **ABSTRACT**

The Water Framework Directive implementation requires a regular ecological status assessment of water bodies within the framework of 6-year management plans. The objective is to avoid the degradation of the ecological status or to restore the Good Ecological State to the deadline foreseen in the SDAGE, *ie* no later than the end of 2027. In that purpose, a first national "Evaluation Circular" was published in July 2005. Since then, this official text has already been subjected to several updates, the previous one (*i.e.* Evaluation Decree of 27 July 2015) still having force of application on most types of French rivers. However, the very large rivers at the meaning of the WFD (catchment area exceeding 10 000 km² at the observation station) were not yet covered by a coherent national assessment system, because of certain specific difficulties encountered in our country and at the European level (in particular, complete absence of true references on such hydrosystems, subjected to significant anthropogenic pressures).

The present action, supported by the Irstea-ONEMA action-sheet N ° 23 "Vegetal bio-indication in watercourses" (2016-2018), aimed to fill this regulatory deficit as quickly as possible by proposing a new national *ad hoc* system harmonized from the start, as far as possible, with the other European Member States. In that aim, the work carried out in the research team CARMA (now ECOVEA) consisted to: on the one hand, gather and analyze comparable national data obtained in a sufficiently homogeneous manner on these types of hydrosystems in the context of surveillance networks (IBD2007 data collected from 2007 to 2013); on the other hand, participate in a European exercise of intercalibration of the evaluation systems of such rivers, whose organization was concomitant to the realization of this national study.

A first intercalibration exercise led by the "Large Rivers" X-GIG was carried out in 2012-2013 on various biological quality elements. In such circumstances, a bit too early compared to the progress of our own national reflection, France participated by providing data and as a mere external observer, but has not been subjected to the intercalibration. During this first exercise, participating countries were able to agree on a consensually defined default level of reference for this type of hydrosystems, and intercalibrated their national methods on that basis. A pressure-impact relationship between the PO₄ content of waters and responses of the intercalibration common metric (diatomic ICM) has been established, from which a shared vision of good status of the LRs has been defined.

Between 2015 and 2017, France took part in the second intercalibration exercise entitled "Fit-in exercise", led under the joint coordination of Bulgaria and Austria, which directly used the results previously obtained by the "LR" X-GIG and involved the participation of 8 different Member States. The results of this second collective exercise, communicated orally and informally during 2017, were validated through the official publication of the Decision EU 2018/229 (February 12th, 2018), authorizing at the same time the enforcement of the new evaluation system proposed by France in the context of that exercise. The incorporation of this new system inside an amendment to the French "Evaluation Decree" has therefore been proposed between October 2017 and February 2018, then signed (July 27th, 2018) and published at the JORF August 28th, 2018). The new assessment system is based on 3 evaluation grids that make it possible to assess in a differentiated and judicious way 3 types of large rivers (poorly-mineralized lowlands; mineralized lowlands; large rivers directly issued from mountain areas).

- **KEY WORDS (THEMATIC AND GEOGRAPHICAL AREA)**

WFD; WATER FRAMEWORK DIRECTIVE; BENTHIC DIATOMS; IBD₂₀₀₇; EVALUATION SYSTEM;
LARGE RIVERS; FRANCE; ECOLOGICAL STATUS



- **SYNTHESE POUR L'ACTION OPERATIONNELLE** [4 pages maximum]

La mise en œuvre de la Directive Cadre sur l'Eau (DCE, JOCE, Octobre 2000) requiert l'évaluation régulière de l'Etat Ecologique des masses d'eau, dans le cadre de plans de gestion de 6 ans. L'objectif poursuivi est d'éviter la dégradation de l'Etat Ecologique, ou de restaurer le Bon Etat Ecologique à l'échéance prévue dans les SDAGE, et au plus tard à fin 2027. Dans cet objectif, une première circulaire nationale « Evaluation » a été publiée en Juillet 2005. Depuis, ce texte a fait l'objet de plusieurs réactualisations sous la forme d'Arrêtés publiés au Journal Officiel de la République Française, dont celui en vigueur au moment de cette étude (Arrêté Evaluation du 27 Juillet 2015) avait force d'application sur la plupart de nos types de cours d'eau nationaux.

Cependant, les très grands cours d'eau (surface cumulée de bassin versant dépassant les 10 000 km² à la station d'observation, au sens de la Commission Européenne) n'y étaient pas encore couverts par un dispositif d'évaluation cohérent, notamment du fait d'une difficulté méthodologique spécifique rencontrée à notre échelle nationale comme à l'échelle Européenne. Le fait qu'il n'existe plus ou pratiquement plus de situations de **référence vraie** sur des systèmes de cette taille qui, au moins à certains endroits de leurs bassins versants, subissent forcément des effets de pressions anthropiques, occasionne une difficulté méthodologique et conceptuelle particulière pour la mesure d'un écart à la **référence adéquate** (cf. DCE). En effet, la procédure classiquement employée par les Etats-Membres sur des hydrosystèmes plus petits, qui se base sur un calcul du niveau de référence moyen ou médian (EQR = 1) calculé, à partir des résultats de la métrique diatomique nationale, sur l'assortiment de sites de référence appartenant au même ensemble biogéographique naturel homogène que les sites à évaluer, n'y est donc pas applicable.

Face à ce type de situation, pour l'évaluation d'état écologique, il est seulement possible de se référer à un **niveau de référence par défaut**, qui correspond à une estimation de l'objectif de niveau d'état écologique atteignable pour de si grands bassins versants en cas de très faible à faible niveau d'influence anthropique (= *least disturbed condition*).

Si chaque Etat-Membre est en capacité légitime de proposer son propre système, basé sur le niveau de référence par défaut qu'il juge adéquat par type national de TGCE, la façon de fixer cet objectif de qualité peut passer par un assortiment varié d'approches bâties sur des procédures méthodologiques très différentes (panel de meilleures notes atteintes, approche de modélisation, voire décision basée sur du simple jugement-expert...). Il est donc particulièrement important que des **exercices d'intercalibration européenne** aient lieu sur ce sujet, à la fois pour dégager une vision collective partagée de l'**objectif de qualité** à atteindre sur de tels types d'hydrosystèmes, mais aussi pour vérifier que les dispositifs nationaux mis au point selon des principes éventuellement très différents conduisent à des **évaluations d'état écologique et d'objectifs de qualité comparables** en inter-Etats-Membres.

Au niveau européen comme dans notre contexte national, historiquement, priorité a tout d'abord été mise sur le calage d'évaluation des hydrosystèmes de plus petite taille. En effet, outre le fait qu'ils constituent une part très dominante des sites de surveillance existants, il s'avérait possible, dans le cadre de leur processus d'évaluation et en déclinant un principe méthodologique relativement établi et partagé : 1) de s'appuyer concrètement sur des situations de référence vraie existantes dans les ensembles naturels concernés, et 2) de construire ainsi des grilles d'EQR basés sur la mesure d'écart par rapport au niveau de référence adéquat.

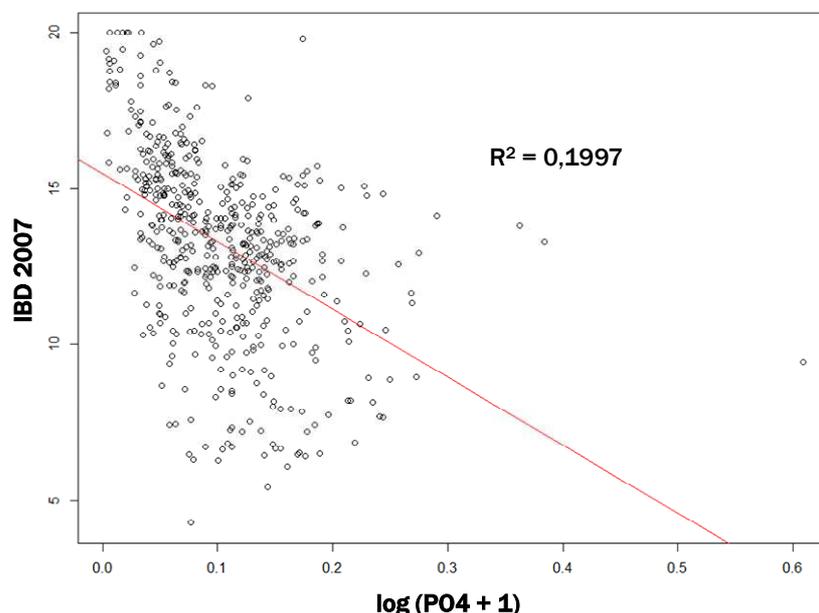
Ces chantiers étant achevés, la présente action, soutenue dans le cadre de la fiche-action Irstea-ONEMA N° 23 « Bio-indication végétale en cours d'eau » (2016-2018), avait pour but de combler au plus vite le déficit réglementaire rencontré vis-à-vis de l'évaluation des TGCE de France, *via* la réalisation d'une étude nationale permettant de proposer un nouveau dispositif dédié à l'évaluation spécifique des hydrosystèmes de ce type, s'appuyant si possible sur une base harmonisée d'emblée avec les autres Etats-Membres Européens.

Dans cette optique, le travail a consisté d'une part à rassembler et à analyser les données comparables, obtenues selon un mode suffisamment homogène sur ces types d'hydrosystèmes dans le cadre de nos réseaux de surveillance nationaux (données IBD₂₀₀₇ recueillies de 2007 à 2013) ; il a d'autre part consisté à participer à un exercice européen d'intercalibration des systèmes d'évaluation de tels cours d'eau, dont l'organisation s'est mise en place dès le début de réalisation de cette étude nationale.

Un 1^{er} exercice d'intercalibration européenne sur ce type de grands hydrosystèmes avait déjà été réalisé en 2012-2013 par le X-GIG "Large Rivers" sur divers maillons-clés biologiques. Dans ce contexte un peu prématuré par rapport aux avancées de notre propre réflexion nationale, la France a participé autant que possible à ce premier exercice, via la fourniture d'un jeu de données national-test et en tant qu'observateur externe, mais ne s'est pas prêtée à l'intercalibration finale. Durant cet exercice, les pays participants ont pu s'accorder sur un niveau de référence par défaut défini sur une base consensuelle, à partir de l'assise du jeu de données d'intercalibration collecté dans le cadre de l'exercice, et ont ensuite inter-calibré leurs méthodes nationales sur cette base. Une relation pression-impact entre la teneur en PO₄ des eaux et les réponses de la métrique commune d'inter-calibration (*i.e.* diatom ICM-bm) a été modélisée, à partir de laquelle les coordinateurs ont décliné une vision partagée du bon état des très grands cours d'eau.

Entre 2015 et 2017, la France a pris part au deuxième exercice d'inter-calibration intitulé "Fit-in Exercise", mené sous la coordination conjointe de la Bulgarie et de l'Autriche avec la participation de 8 États membres différents. La coordinatrice de l'exercice a sollicité auprès des États-Membres les principaux éléments de description des systèmes nationaux d'évaluation présentés et demandait aussi, autant que possible, d'apporter les éléments illustratifs de leur capacité à évaluer le type de pression anthropique sur lequel l'exercice était réalisé (ici, PO₄), sur le plan général mais aussi plus spécifiquement sur TGCE.

Comme tous les autres pays-Membres, la France illustre une relation pression-impact atténuée entre son outil national (IBD₂₀₀₇) et le gradient de PO₄ en TGCE, avec un R² se réduisant de 0,457 (cf. Figure 11 du rapport) sur un jeu de données « tous cours d'eau », à environ 0,2 strictement sur TGCE (cf. ci-dessous).



Relation pression-impact obtenue entre le gradient d'orthophosphates et l'IBD₂₀₀₇ sur l'assise nationale de données TGCE (seuil de taille : 8 000 km²)

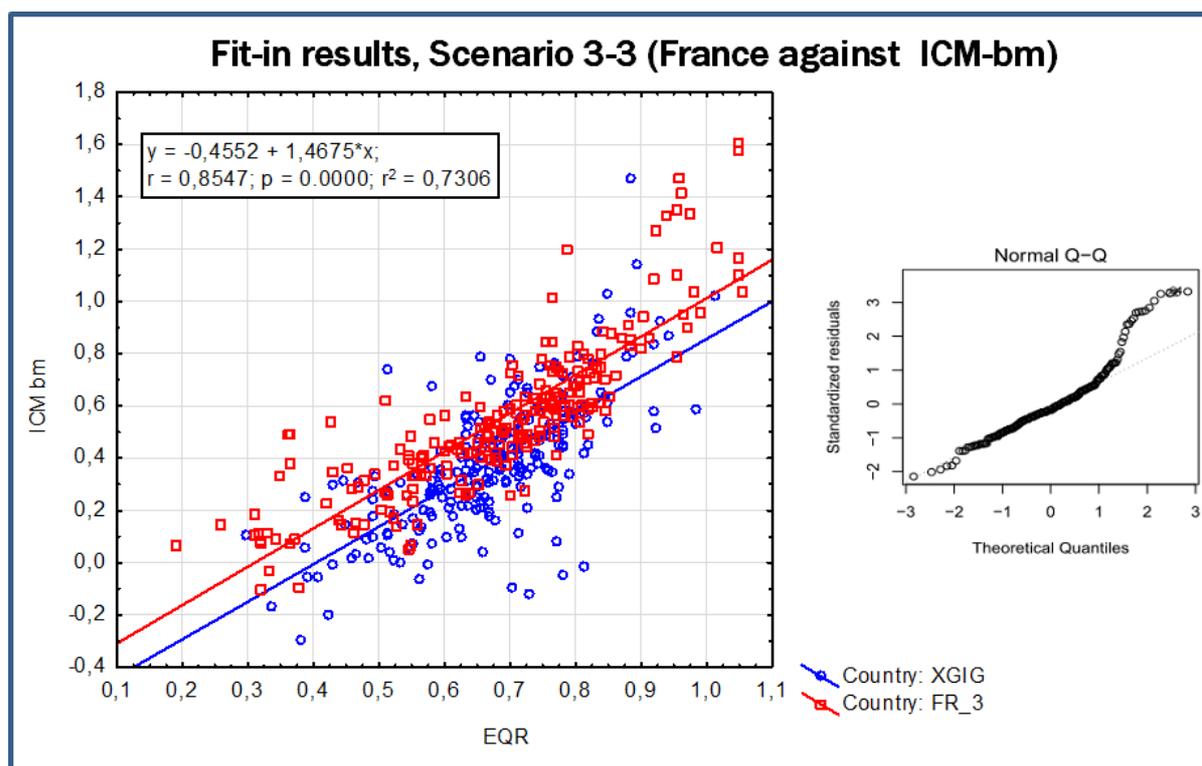
Le graphique présenté ci-dessus donne l'occasion d'évoquer 2 aspects pouvant affaiblir la qualité de cette relation pression-impact mais aussi, quelque part, la robustesse des résultats de l'exercice d'intercalibration réalisé. En effet, par rapport à la physionomie du nuage de points, l'ajustement linéaire pression-impact obtenu, avec un niveau élevé de résidus, suggère l'influence de 2 types de problèmes :

- 1) Dans certaines situations (quartier en bas à gauche de l'ajustement linéaire obtenu), l'abaissement de note d'IBD₂₀₀₇ observé ne semble pas être le reflet de la pression de PO₄ telle qu'analysée au site, mais pourrait être généré par un ou plusieurs autres paramètres de pression anthropique venant se cumuler ou se substituer aux effets du seul phosphore. L'IBD ayant été calé sur la réponse des diatomées à un gradient intégré de pression basé sur 7 paramètres différents de pression anthropique, il n'est pas anormal que ces 7 paramètres participent conjointement au degré d'altération biologique mesuré au site, ni d'obtenir une relation relativement floutée vis-à-vis d'un gradient monoparamétrique tel que présenté ici.
- 2) Une autre hypothèse suggérée à la fois par la physionomie de cet ajustement et par la perte générale de qualité de la relation établie sur un jeu de données spécifique « TGCE », par rapport à une assise de cours d'eau « tous types confondus », serait que la représentativité temporelle d'une stratégie basée sur des échantillonnages ponctuels de chimie, déjà peu satisfaisante sur de plus petits hydrosystèmes, serait encore moins représentative de la qualité intégrée de la colonne d'eau dans le cas des Très Grands Cours d'Eau, contribuant ainsi à flouter un peu plus la qualité de la relation établie.

Pour pouvoir participer à l'exercice collectif dit « de Fit-in », il fallait au préalable, dans des conditions d'échéancier très tendu, mettre au point et proposer un prototype de dispositif national d'évaluation des TGCE, dont l'application et les résultats puissent servir de base concrète pour la réalisation du processus d'inter-comparaison avec les résultats de l'exercice précédent (« LR » X-GIG).

Ainsi, entre Décembre 2015 et le 16-06-2016, deadline pour permettre le bouclage du rapport du 2^{ème} exercice d'intercalibration, ce sont successivement 3 prototypes de dispositif national qui ont fait l'objet de tests préalables, suivis par la soumission d'un dernier essai de tunage fin du 3^{ème} scénario testé.

Dans le cadre de la méthodologie d'inter-comparaison prévue pour la réalisation de cet exercice collectif, l'application du Scénario 3 - Variante 3 finalement choisi en accord avec les tutelles a procuré les résultats présentés ci-dessous (points rouges et ajustement rouge pour la France), en comparaison avec les résultats de l'exercice X-GIG précédent (figurés en couleur bleue).



En fonction du cadre interprétatif proposé au niveau européen (cf. Guidance Ecostat N° 30) et malgré les ajustements à la baisse réalisés au cours des itérations précédentes, l'ajustement linéaire obtenu sur la base des EQR-IBD₂₀₀₇ Français présente encore un décalage vers le haut par rapport aux résultats finaux de l'exercice précédent, générant des biais positifs interprétés ici comme l'expression d'un système d'évaluation Français un peu plus sévère, sur ce type européen « Large Rivers », que la vision collective du Bon Etat. Cependant, sans qu'il n'y ait d'intention d'évaluer nos TGCE de façon plus rigoureuse que ceux de nos Pays-Membres voisins, plusieurs éléments interviennent en modération de ces conclusions, notamment :

- Evaluation par l'IBD de 7 types d'altérations (et non de l'unique altération occasionnée par le PO4), ce qui conduit à une évaluation d'altération biotique intégrée d'un niveau un peu supérieur en moyenne,
- Assise initiale d'apprentissage de l'étalon commun modélisé (ICM-bm) comportant une lacune en situations de haute qualité naturelle rencontrées pendant le 2^{ème} exercice (France et Italie notamment), ce qui interroge sur la validité, sur le plan biogéographique, de l'apprentissage et du calage de cet étalon réalisé pendant le 1^{er} cycle, pour évaluer de façon satisfaisante l'intégralité du 2^{ème},
- Méthodologie d'intercalibration entre métrique nationale et étalon commun modélisé (ICM-bm) basée sur une intercomparaison d'ajustements linéaires alors que, dans le cas de la France, la relation entre ces 2 métriques est manifestement curvilinéaire (cf. à droite physionomie déséquilibrée des résidus).

Quoi qu'il en soit, l'enjeu essentiel de cet exercice était de **pouvoir intercalibrer de façon valide le prototype Français soumis au processus** collectif d'intercalibration, en entrant dans la fourchette de biais jugés tolérables dans le cadre de l'exercice (+/- 0,5). Le tableau suivant montre que c'est bien le cas pour le scénario finalement choisi pour la France (dernière ligne, scénario 3-FR, variante 3-3, voir biais calculés et décision en bout de ligne à droite).

National specificities			National boundaries EQR					Boundaries on ICM + offset				BIAS		Result	
Country - Scenario	Offsets	*national type	H/G	G/M	M/P	REF	P/B	REF	H/G on ICM	G/M on ICM	M/P on ICM	H/G bias	G/M bias	H/G	G/M
FR - Scenario 1	0,2047	IBD-EQR (every type)	0,94	0,78	0,55	1	0,3	0,8019	0,7154	0,4846	0,1528	0,4868	0,5201	passed	passed
FR3 - Scenario 3-1	0,2032	IBD-EQR (every type)	0,92	0,77	0,52	1	0,26	0,8091	0,6917	0,4716	0,1047	0,40295	0,43495	passed	passed
FR3 - Scenario 3-2	0,2032	IBD-EQR (every type)	0,92	0,765	0,52	1	0,26	0,8091	0,6917	0,4642	0,1047	0,38995	0,42342	passed	passed
FR3 - Scenario 3-3	0,2032	IBD-EQR (every type)	0,92	0,76	0,52	1	0,26	0,8091	0,6917	0,4569	0,1047	0,3778	0,4114	passed	passed

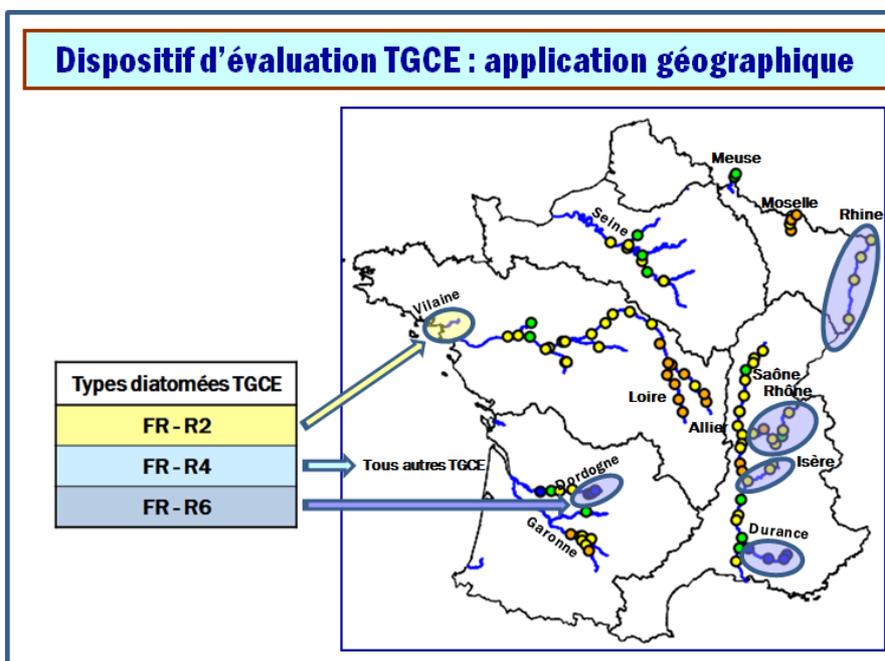
* adjustment strongly recommended

Le calage de notre nouveau système national d'évaluation des TGCE (*i.e.* Scénario 3-3 FR) a été inter-calibré avec succès (obtention d'un visa « OK »), et a donc jugé DCE-compatible (cf. Décision d'Intercalibration U.E. du 12-02-2018). Les grilles sur lesquelles se base ce dispositif sont présentées ci-dessous :

Dispositif d'évaluation TGCE intercalibré (Scénario 3-3 FR du 16-06-2016)

IBD ₂₀₀₇	REF	H/G	G/M	M/P	P/B	MIN
EQR-IBD ₂₀₀₇	1	0,92	0,76	0,52	0,26	0
FR - R2	17,4	16,088	13,464	9,528	5,264	1,0
FR - R4	18,1	16,732	13,996	9,892	5,446	1,0
FR - R6	19,1	17,652	14,756	10,412	5,706	1,0

L'intégration de ce nouveau système d'évaluation dans un modificatif de l'Arrêté "Evaluation" français a été proposée entre octobre 2017 et février 2018. Ce nouvel Arrêté, devenu officiel depuis sa publication (signature le 27/07/2018, parution au JORF du 28-08-2018), est entré en vigueur au 01/01/2019. Les tableaux édités présentaient quelques erreurs accidentellement introduites par le service chargé du maquetage du JORF. Un rectificatif devrait être publié très prochainement.



Ce nouveau système d'évaluation se base sur une nouvelle grille en EQR-IBD₂₀₀₇ spécifiquement calée pour les TGCE (voir tableau un peu plus haut), aux seuils de qualité un peu rabaisés par rapport à celle qui s'applique sur les autres cours d'eau de notre territoire national.

En fonction des types diatomiques présents et des niveaux de référence par défaut à appliquer par grand ensemble naturel, cette grille en EQR se décline en 3 grilles spécifiquement adaptées pour évaluer de manière différenciée nos 3 types de grands fleuves nationaux, à savoir : **Type 2** : plaines peu minéralisées ; **Type 4** : plaines minéralisées ; **Type 6** : grands cours d'eau directement issus de massifs montagneux (cf. figure ci-dessus).

SOMMAIRE

1.	INTRODUCTION	1
2.	BILAN DES DONNEES NATIONALES TGCE ET TRAVAIL SUR DONNEES	3
2.1.	Aspects généraux, éléments de contexte :	3
2.2.	Aspects particuliers, difficultés à résoudre :	4
2.3.	Etat des données nationales TGCE / diatomées :	6
2.3.1.	Historique et portée des requêtes :	6
2.3.2.	Données de description générale des sites (stations) :	6
2.3.3.	Données biologiques (données diatomiques) :	8
2.3.4.	Croisement avec les données de qualité d'eau, bilan quantitatif :	9
2.4.	: Post-traitement et complétion des données abiotiques	12
2.4.1.	Données de qualité d'eau (aspects généraux).....	12
2.4.2.	Gestion du problème d'absence de données numériques réelles.....	13
2.4.3.	Réalisation de l'agrégation des données disponibles	14
2.5.	Post-traitement et complétion des données hydrochimiques spécifiques	15
2.5.1.	Aspects généraux.....	15
2.5.2.	Régression TAC / carbonates totaux	16
2.5.3.	Régression TAC / Ca++	17
2.5.4.	Régression TAC / [Cond. El. * pH].....	19
2.6.	Calculs d'alcalinité au site	22
2.7.	Bilan des données post-traitées, sélection du jeu national envoyé au GIG	23
3.	ETUDE DES RELATIONS PRESSION-IMPACT BASEES SUR L'IBD ₂₀₀₇	25
3.1.	Aspects généraux, éléments de contexte	25
3.2.	Comparaison des réponses de l'IBD ₂₀₀₇ et de l'IPS :	25
3.3.	Relations pression-impact entre IBD ₂₀₀₇ et nutriments (tous types de cours d'eau confondus).....	26
3.3.1.	IBD et PO₄ :	26
3.3.2.	IBD et NH₄ :	27
3.4.	Relations pression-impact entre IBD ₂₀₀₇ et nutriments sur TGCE	27
3.4.1.	IBD₂₀₀₇ et PO₄ (analyses 2015, moyenne annuelle).....	28
3.4.2.	IBD₂₀₀₇ et NH₄ (analyses 2015, moyenne annuelle)	32
3.4.3.	IBD₂₀₀₇ et nutriments (analyses 2017, comparaison des temps d'intégration).....	33
3.4.4.	IBD₂₀₀₇ et nutriments (requêtes 2017, relations consolidées)	38
3.5.	Relation pression-impact sur la base d'un gradient composite TGCE	40
4.	Liens entre IBD ₂₀₀₇ et attributs autoécologiques des communautés	43
4.1.	Liens autoécologiques IBD ₂₀₀₇ / trophie (PO ₄), tous types de cours d'eau confondus	44
4.2.	Liens IBD ₂₀₀₇ / trophie sur une sélection de Très Grands Cours d'Eau	45
4.3.	Liens entre les notes d'IBD ₂₀₀₇ et les caractéristiques de saprobie sur une sélection de Très Grands Cours d'Eau.....	49
4.4.	Conclusions sur les relations pression-impact de l'IBD ₂₀₀₇ sur TGCE.....	52

5.	PROPOSITION ET STABILISATION D'UN SYSTEME FRANCAIS D'EVALUATION DES TGCE	53
5.1.	Eléments généraux et de contexte	53
5.2.	Dispositif d'évaluation en vigueur, répercussions vis-à-vis des TGCE.....	55
5.3.	Examen des données existantes et repérage des cas particuliers	59
5.4.	Premier prototype de dispositif Français TGCE proposé	62
5.4.1.	Liaison avec les échanciers de travail du GIG Large Rivers	62
5.4.2.	Réadaptation du dispositif national au cas des TGCE : 1^{ère} proposition.....	62
5.4.3.	Résolution des cas particuliers et 1^{er} dispositif d'évaluation proposé :	64
5.4.4.	Résumé de l'exercice d'intercalibration N° 1 du X-GIG « Large Rivers »	70
5.4.5.	1^{ère} itération de l'exercice d'intercalibration 2015-2016 « Fit-in exercise »	73
5.4.6.	2^{ème} itération du processus d'intercalibration (V2 du prototype d'évaluation des TGCE de France)	83
5.4.7.	3^{ème} itération du processus d'intercalibration (V3 du prototype d'évaluation des TGCE de France)	87
6.	APPUI A LA MISE EN OEUVRE DU NOUVEAU SYSTEME NATIONAL D'EVALUATION DES TGCE.....	95
6.1.	Critères de définition des TGCE et étude des conséquences sur l'évaluation des TGCE de France	96
6.1.1.	Aspects généraux et contextuels.....	96
6.1.2.	Descriptif de 4 scénarios envisageables pour la mise en œuvre de l'évaluation des TGCE	99
6.1.3.	Assise et résultats bruts des 4 scénarios d'évaluation	100
6.1.4.	Inter-comparaison et discussion des résultats des 4 scénarios	104
6.2.	Proposition d'évolution du contenu de l'Arrêté Evaluation 2018.....	108
6.2.1.	Premières propositions d'évolution du contenu de l'Arrêté Evaluation :	108
6.2.2.	Nouvelle phase de correction de l'Arrêté Evaluation du 27 Juillet 2018 :	111
7.	CONCLUSIONS.....	115
8.	SIGLES & ABREVIATIONS.....	117
9.	BIBLIOGRAPHIE	119
	ANNEXE 1 : Evaluation d'état diatomique des cours d'eau candidats-TGCE de France (au seuil de 8 000 km ² de BV intégré)	121
10.	REMERCIEMENTS	131

Table des Figures

Figure 1 : Problèmes rencontrés pour la mise en relation entre relevés diatomiques TGCE et données abiotiques de chimie (requêtes PANDORE 2015)	9
Figure 2 : Mise en relation entre relevés diatomiques et données abiotiques de chimie pour l'étude nationale TGCE (requêtes PANDORE 2017).....	10
Figure 3 : Régression linéaire TAC /Carbonates totaux (255 doublets de données existantes)	16
Figure 4 : Reconstitution de données manquantes à l'aide de la régression TAC /Carbonates totaux.....	17
Figure 5 : Relation entre calcium et TAC (255 doublets de données existantes)	18
Figure 6 : Régression linéaire tout-venant entre (pH * Cond. El.) et TAC	19
Figure 7 : Régression linéaire entre (pH * Cond. El.) et TAC.....	20
Figure 8 : Régression linéaire consolidée entre [pH * Cond. El.] et TAC.....	21
Figure 9 : Jeu de données complété sur la base de la relation [TAC = 0,0041 (pH * Cond.El)].....	22
Figure 10 : Régressions linéaires entre les 2 versions successives de l'IBD et l'IPS, indice diatomique de référence	25
Figure 11 : Relation pression-impact entre les 2 versions successives d'IBD et le log (PO ₄ +1)	26
Figure 12 : Pressure-impact relationships between the 2 successive versions of IBD and log (NH ₄ + 1)...	27
Figure 13 : Relation linéaire entre IBD ₂₀₀₇ et log (PO ₄ + 1) (moyenne annuelle) sur le jeu de données « Très Grands Cours d'Eau	28
Figure 14 : Relation linéaire entre IBD ₂₀₀₇ et log (PO ₄ + 1) (moyenne annuelle) en Très Grands Cours d'Eau après enlèvement d'1% d'outliers	29
Figure 15 : Relation linéaire entre les notes d'IBD 2007 et le log (NH ₄ + 1) (moyenne annuelle) dans les Très Grands Cours d'Eau	32
Figure 16 : Relations pression-impact établies entre IBD et nutriments sur TGCE de France (requête PANDORE Juillet 2017)	35
Figure 17 : Relation pression-impact entre IBD et PO ₄ sur TGCE (intégration temporelle de 90 jours)	38
Figure 18 : Relation pression-impact entre IBD et NH ₄ sur TGCE (intégration temporelle de 90 jours)	39
Figure 19 : ACP sur 7 descripteurs d'anthropisation, utilisée pour la genèse de l'IBD ₂₀₀₇ (partie gauche) ; détermination des 7 classes de qualité correspondantes (partie droite).....	40
Figure 20 : ACP réalisée à partir des 7 descripteurs d'anthropisation de l'IBD ₂₀₀₇ , sur jeu de relevés nationaux strictement TGCE (partie gauche) ; détermination des 7 classes de qualité correspondantes (partie droite).	41
Figure 21 : Relation pression-impact établie sur TGCE entre scores de gradient composite et notes d'IBD ₂₀₀₇	42
Figure 22 : Caractéristiques trophiques des communautés diatomiques de France sur sites de référence et au long du gradient de classes de qualité des eaux (depuis le très bon état jusqu'au mauvais état)	44
Figure 23 : Pourcentage d'espèces sensibles au PO ₄ -par classe de qualité en TGCE (> 8 000 km ²).....	45
Figure 24 : Distribution des espèces de la Classe 7 (mal caractérisées vis-à-vis de la trophie) au long du gradient des 5 classes de qualité des Très Grands Cours d'Eau	47
Figure 25 : Pourcentage d'espèces tolérantes au PO ₄ par classe de qualité en TGCE (> 8 000 km ²).....	48
Figure 26 : Pourcentage d'espèces sensibles à l'altération saprobique par classe de qualité dans les TGCE (> 8 000 km ²)	49
Figure 27 : Distribution des espèces au statut indéterminé vis-à-vis de la saprobie au long du gradient des 5 classes de qualité	50

Figure 28 : Pourcentage d'espèces tolérantes à la saprobie par classe de qualité dans les TGCE (> 8 000 km ²).....	51
Figure 29 : Zonage des régions diatomiques naturelles en vigueur pour l'évaluation d'état écologique (Version 2 de 2012 incluse dans l'Arrêté Evaluation du 27 Juillet 2015).....	55
Figure 30 : Grilles d'évaluation diatomique actuelles (Arrêté Evaluation du 27-07-2015).....	56
Figure 31 : Grille d'évaluation diatomique des TGCE actuellement en vigueur ('Arrêté du 27-07-2015)	57
Figure 32 : Bilan des 127 sites nationaux TGCE associant descripteurs abiotiques chimiques et relevés diatomiques.....	59
Figure 33 : Classification automatique de la grille d'évaluation selon la Région Diatomées d'appartenance ; illustration des cas spécifiques à expertiser (en rouge).....	60
Figure 34 : 1 ^{ère} proposition de grilles d'évaluation de l'Etat Ecologique pour les Très Grands Cours d'Eau : a) en notes d'IBD 2007 (arrondies au 1/1000 ^{ème}) ; b) en grille unique d'EQRs.....	63
Figure 35 : 1 ^{er} dispositif d'évaluation proposé pour les TGCE du bassin hydrographique du Rhône.....	66
Figure 36 : 1 ^{er} dispositif d'évaluation proposé pour les TGCE du Nord-Est de la France.....	67
Figure 37 : 1 ^{er} dispositif d'évaluation proposé pour les TGCE de l'Ouest de la France.....	68
Figure 38 : 1 ^{er} dispositif d'évaluation proposé pour l'hydrosystème Garonne-Dordogne.....	69
Figure 39 : les 2 approches fondamentales du benchmarking en intercalibration : a) approche se référant à des sites de référence et à un gradient de pression comparable ; b) » Approche de benchmarking continu (données des pays couvrant des gradients discontinus).....	71
Figure 40 : Réajustement des EQR nationaux (rouge et vert) en fonction du décalage systématique mesuré par rapport à l'ajustement modélisé (en noir) de l'ICM (modèle linéaire à effets mixtes).....	72
Figure 41 : Organigramme de la méthodologie suivie par le X-GIG pour le 1 ^{er} exercice d'intercalibration (2011-2012).....	73
Figure 42 : Gradients de PO ₄ transformés Ln (à gauche) et réponses diatomiques en EQR nationaux (à droite) présentés par les jeux de données du Fit-in Exercice « large Rivers » 2015 – 2016.....	75
Figure 43 : Relation entre IPS observé et IPS benchmark prédit (en fonction de la relation avec le PO ₄ pré-établie dans le cadre du précédent X-GIG "Large Rivers").....	77
Figure 44 : 1 ^{ère} mise en relation entre EQRs nationaux d'évaluation diatomique du 2 nd Fit-in exercice, du jeu de données du 1 ^{er} X-GIG exercice (points et ajustement en bleu), avec l'ICM-bm modélisé.....	78
Figure 45 : Régression linéaire entre la méthode d'évaluation française (Scénario 1) et l'ICM-bm modélisé, en inter-comparaison avec le jeu de données du précédent X-GIG.....	80
Figure 46 : Répartition des résidus en liaison avec l'ajustement linéaire produit pour la France en Figure 45.....	81
Figure 47 : Relation entre les EQR-IBD Français par type de TGCE et l'ICM-bm (1 ^{er} Scénario).....	82
Figure 48 : Test d'intercalibration du 2 ^{ème} scénario d'évaluation des TGCE de France (30/05/2016).....	84
Figure 49 : Régression linéaire entre la méthode d'évaluation française (Scénario 2-2) et l'ICM-bm modélisé, en inter-comparaison avec le jeu de données du précédent X-GIG.....	85
Figure 50 : Biais d'intercalibration calculés pour le Scénario 2-2 de dispositif Français d'évaluation TGCE.....	86
Figure 51 : Impact d'un changement de grille d'évaluation pour les TGCE du Type 5 (Alpin).....	88
Figure 52 : Grille d'évaluation des TGCE de France selon le Scénario 3 (proposition du 03-06-2016).....	89
Figure 53 : Dispositif Français, Scénario 3 : Evolution des biais par rapport au Scénario 1.....	90
Figure 54 : Essai de tunage fin de grilles d'EQR-IBD2007 dérivées du scénario 3 (06 au 13-06-2016).....	90
Figure 55 : Calcul de biais sur les 3 variantes du Scénario 3-FR (13-06-2016).....	91
Figure 56 : Grille finale d'évaluation des TGCE proposée par la France (Scénario 3-3, 16-06-2016).....	93
Figure 57 : Dispositif final d'évaluation TGCE à appliquer sur le bassin du Rhône (16-06-2016).....	94
Figure 58 : Résultats comparés d'évaluation au relevé procurés par les 4 scénarios étudiés à l'échelle du jeu de données global des 671 relevés diatomiques.....	104

Table des Tableaux

Tableau 1 : Description quantitative globale des données nationales disponibles selon la date de requête et le type d'utilisation (étude nationale versus Intercalibration Européenne)	10
Tableau 2 : Nombre moyen de valeurs de chimie impliquées dans le calcul des valeurs moyennes	34
Tableau 3 : Colonne centrale : pondération appliquée aux paramètres IBD (selon niveau de contribution calculé dans l'IPS); à droite : poids relatif issu de l'ACP pour le calcul du gradient composite d'altération.....	41
Tableau 4 : Jeux de données nationaux complets finalement fournis par les Etats-membres participants	74
Tableau 5 : Résultats d'évaluation diatomique au relevé selon classification TGCE basée sur la typologie nationale : Scénario 1 : Situation actuelle (cf. Arrêté du 27-07-2015 actuellement en vigueur).....	101
Tableau 6 : Résultats d'évaluation diatomique au relevé des TGCE définis selon la typologie nationale : Scénario 2 : Application des nouvelles grilles 2018 sur TGCE, règles 2015 sur cours d'eau non-TGCE.....	102
Tableau 7 : Résultats d'évaluation diatomique au relevé selon le Scénario 3 (application des nouvelles grilles d'évaluation sur les TGCE définis au seuil de surface > 8 000 km ²).	103
Tableau 8 : Résultats d'évaluation diatomique au relevé selon le Scénario 4 (application des nouvelles grilles d'évaluation sur les TGCE définis au seuil de surface > 10 000 km ² . Grilles 2015 sur cours d'eau plus petits).....	104

Elaboration d'un dispositif d'évaluation des Très Grands Cours d'Eau (TGCE) de France et intercalibration Européenne : Rapport Final

1. INTRODUCTION

La mise en œuvre de la Directive Cadre sur l'Eau (REF 1) implique l'Evaluation de l'Etat Ecologique des masses d'eau dans le cadre de plans de gestion de 6 ans. Dans le cas où leur état est qualifié de très bon ou bon, elles doivent se maintenir dans leur état écologique actuel sans dégradation supplémentaire. Pour celles qui ont été évaluées en-dessous du Bon Etat, la DCE requiert de leur faire regagner le Bon Etat Ecologique aussitôt que possible (et au plus tard en 2030), au moyen de mesures correctives mises en œuvre dans le cadre de plans de gestion adaptés (PGME).

Sur cours d'eau, concernant le maillon biologique des diatomées benthiques, un dispositif national d'évaluation DCE-compatible, qui s'applique sur la plupart des cours d'eau permanents, a été mis en place depuis Juillet 2005. Depuis cette date de début de mise en œuvre, il a subi 2 évolutions et adaptations notables dont celle actuellement applicable pour la 1^{ère} moitié du Plan de Gestion des Masses d'Eau en cours (PGME 2016-2021), qui est entrée en vigueur via la publication officielle de l'Arrêté Evaluation actuel signé en date du 27 Juillet 2015, publié au J.O. de la République Française du 27 Août 2015 (REF 2). Dans le même temps, des grilles nationales d'évaluation de l'Etat Ecologique destinées à s'appliquer sur les petits, moyens et grands cours d'eau de notre territoire national ont été inter-calibrées avec succès, au fur et à mesure de leur évolution, dans le cadre de 3 groupes Européens d'inter-calibration (GIG Central-Baltique, GIG Alpin, GIG Méditerranéen) et de 2 rounds successifs d'intercalibration, ces exercices ayant été menés selon une logique prenant en compte la situation géographique et la typologie européenne des cours d'eau concernés.

Les très grands cours d'eau, qui présentent une surface cumulée de bassin versant de grande taille (> 10 000 km² à la station d'observation, selon la typologie DCE « Large Rivers »), posent un problème spécifique qui a retardé la mise au point des systèmes d'évaluation au niveau national et européen. En effet, leur taille minimale correspond à des stations de suivi implantées en position relativement aval de grands bassins versants (fleuves principaux, affluents majeurs avant leur confluence avec le drain principal), où l'anthropisation est déjà omniprésente. De plus, beaucoup d'entre eux ont fait l'objet d'aménagements mettant en jeu des dispositifs physiques plus ou moins importants. La conséquence est qu'il n'existe plus ou pratiquement plus, à l'échelle de la France aussi que, plus largement, à l'échelle de l'Europe de l'Ouest, de situation(s) de référence vraie permettant de caler facilement les dispositifs d'évaluation. Dans ce contexte, il n'est pas possible de décliner vers l'aval la même logique de construction de dispositif que celle utilisée sur des cours d'eau plus petits (évaluation de l'atteinte à l'état écologique établie par écart à la situation de référence adéquate, via le calcul d'EQR). Cette évaluation doit donc se baser sur un **niveau de référence par défaut (= least disturbed condition)**, défini autant que possible sur une base consensuelle par type européen de Très Grand Cours d'Eau, puis par une déclinaison en classes de qualité procurant une évaluation cohérente d'un Pays-Membre à l'autre.

Entre 2011 et 2012, un premier groupe européen d'intercalibration doté de jeux de données de 13 Etats-Membres (le X-GIG « Large Rivers ») avait réalisé un premier exercice d'intercalibration de systèmes nationaux d'évaluation basés sur différents maillons biologiques des TGCE, dont les macro-invertébrés et les diatomées benthiques (REF 3). A l'instar de 8 autres pays-membres, la France, qui ne disposait pas encore du recul-données suffisant ni d'un dispositif spécifique pour l'évaluation de ses très grands cours d'eau, a participé à l'exercice à titre d'observateur externe et a fourni un jeu de données national-test, mais n'était pas en mesure de faire intercaler un dispositif d'évaluation national qui n'était pas encore créé. A cette occasion, seuls 5 pays-membres ont pu faire évaluer avec succès leur dispositif national d'évaluation « Large Rivers » par le compartiment diatomique, ces très grands cours d'eau ayant été ventilés dans 2 types européens d'intercalibration en fonction de leur géochimie-hydrochimie naturelle (eaux peu minéralisées, Alk < 0,5 meq CO3/l versus eaux moyennement à fortement minéralisées, Alk > 0,5 meq CO3/L).

Dans notre contexte national, dans un premier temps, priorité avait été donnée à l'évaluation des cours d'eau de France de petite taille (quelques km²) à grande taille (< 10 000 km² de surface intégrée de bassin versant), de loin les plus nombreux, qui répondaient à un schéma classique d'évaluation bien décrit dans des documents-guides ECOSTAT. En effet, la déviation par rapport au niveau de référence s'établissait à partir d'un calcul assez simple d'EQR (équivalent de qualité écologique), calé sur un niveau de référence vraie (=1) calculé sur la zone géographique naturelle (ou HER) concernée, et pouvant varier entre 1 et 0 selon le niveau d'impact anthropique subi.

L'évaluation des cours d'eau Français répondant à ce schéma le plus classique ayant été validée comme DCE-compatible entre 2007 et 2012 (visa OK de l'U.E. obtenu dans le cadre des inter-comparaisons européennes menées dans différents GIGs), la poursuite de la mise en application de cette Directive passait désormais par la fourniture de systèmes nationaux d'évaluation dédiés à des types particuliers d'hydrosystèmes posant plus de problèmes méthodologiques. Ainsi, la finalisation de la mise au point d'un système d'évaluation national spécifiquement destiné aux TGCE a été inscrite au programme de la nouvelle fiche-action AFB-Irstea N° 23 «Bio-indication végétale en cours d'eau» (programmation trisannuelle 2016-2018).

Suite au premier chantier d'intercalibration des très grands cours d'eau mené par le X-GIG « Large Rivers », dont la réalisation a été présentée sommairement un peu plus haut, l'annonce du démarrage d'un deuxième exercice d'intercalibration Européen de ce type, dont l'organisation s'est mise en place rapidement à partir de la mi-2015 sous une coordination conjointe de la Bulgarie et de l'Autriche, a ouvert l'opportunité rare de pouvoir soumettre directement les résultats d'évaluation d'un prototype de dispositif d'évaluation national des TGCE à l'intercalibration européenne, puis de l'actualiser au fur et à mesure si nécessaire afin qu'il soit d'emblée correctement calé sur la pratique d'évaluation faisant référence au niveau Européen.

En représentation de notre collectivité nationale, l'équipe CARMA d'Irstea Bordeaux s'est donc impliquée activement dans cet exercice collectif qui a débuté en début de 2^{ème} semestre 2015, soit en avance d'environ un semestre par rapport à l'entrée en vigueur effective de la fiche-action précitée. En effet, la participation à cet exercice impliquait de répondre avec une bonne réactivité aux diverses demandes du GIG nécessaires à la bonne réalisation de l'exercice, à savoir :

- fourniture du jeu de données national et de métadonnées dans des délais contraints,
- fourniture, via une note d'expertise nationale (**REF. 4** : Guéguen et al, 2015), de signalétiques des outils et méthodes d'évaluation utilisées à notre niveau national, des formules de calcul,
- démonstration, via une note d'expertise nationale (**REF. 5** : Guéguen et al, 2015), de relations pressions-impacts illustratives des capacités d'évaluation de notre outil national basé sur l'indice diatomique « IBD₂₀₀₇ », tous cours d'eau confondus et, si possible, plus spécifiquement sur TGCE,
- signalement au fil de l'eau des problèmes ou des incompréhensions rencontrées vis-à-vis du déroulement de l'exercice,
- réponse rapide aux premières tentatives d'intercalibration réalisées et modification en conséquence de notre prototype de système national, afin de rentrer dans la fourchette d'écart jugés acceptables.

Par ailleurs, notre équipe de recherche a dû procéder à des requêtes permettant de fournir l'assise actualisée maximale à l'étude nationale prévue dans la Fiche-Action, qui visait à analyser en profondeur tous les référentiels nationaux actualisés disponibles et à proposer le dispositif le plus adéquat d'évaluation des TGCE de France dans un cadre moins contraint dans le temps. Ces 2 chantiers se sont donc un peu télescopés du fait que les demandes du Groupe d'Intercalibration, qui intervenaient de façon anticipée par rapport à notre propre planification de travail, portaient sur une assise-données proche de celle concernant la réalisation de l'étude nationale, mais imposaient dans le même temps l'acquisition d'un recul suffisant sur les référentiels et sur la pertinence du système d'évaluation proposé dans notre contexte naturel et typologique national.

Le présent document constitue le rapport final de cette étude TGCE. Il récapitule les éléments disponibles pour la réalisation de l'**étude nationale** et les démarches et résultats capitalisés dans ce cadre. L'assise de données mobilisée n'ayant pas été exactement la même, il récapitule aussi de façon différenciée les informations et les démarches réalisées en relation avec le 2^{ème} **exercice d'intercalibration européenne « Fit-in Large Rivers »**.

Dans le cadre de cet exercice collectif, l'ajustement du niveau d'évaluation Français, comme celui de plusieurs autres pays, a nécessité plusieurs itérations avant de rentrer dans les biais acceptables par rapport à la vision dominante du Bon Etat telle qu'elle avait été établie par le précédent X-GIG « Large Rivers » durant le 1^{er} exercice d'intercalibration (2011-2012). La dernière adaptation du prototype de dispositif Français TGCE, proposée en date du 17-06-2016, a été évaluée pleinement DCE-compatible. C'est donc cette version, reportée en fin de rapport, qui a servi de base à la rédaction d'un contenu d'un modificatif d'Arrêté Evaluation signé le 27-07-2018, puis publié au Journal Officiel de la République Française. Ce texte officialise le nouveau dispositif d'évaluation qui aura force d'application pour la 2^{ème} moitié du PGME en cours (2019-2021).

En complément, l'équipe de recherche ECOVEA a participé courant 2018 à une adaptation-réédition du guide « Evaluation » intégrant les nouvelles modifications du dispositif national, ce document technique ayant pour but d'appuyer le passage du nouveau dispositif dans une application de routine. Au 1^{er} semestre 2019 et dans un cadre contractuel différent (ce chantier relevant plutôt d'une fiche-action Aquaref-S3E), un appui complémentaire sera encore à fournir à l'AFB en vue d'une adaptation du S3E visant à intégrer ce complément de dispositif.

2. BILAN DES DONNEES NATIONALES TGCE ET TRAVAIL SUR DONNEES

2.1. Aspects généraux, éléments de contexte :

L'inventaire des données nationales disponibles, via des requêtes dans la base PANDORE, s'est confronté en premier lieu à un problème de **définition de ce que sont des Très grands Cours d'Eau**. En effet, différents critères peuvent contribuer à définir ces objets et leur taille, comme la distance à la source au point d'observation, la surface cumulée de bassin versant drainé au point d'observation, la largeur, la section mouillée, le débit nominal au point d'observation, l'ordre de Strahler... Tous ces critères, pris séparément ou considérés de façon plus intégrée, peuvent être utilisables pour représenter un gradient de taille des hydrosystèmes « cours d'eau » et pour en établir une typologie basée sur leur importance.

La typologie nationale sous-jacente à l'application des différents Arrêtés de Surveillance et d'Evaluation parus depuis 2005, année de mise en place du premier dispositif national de surveillance dans le cadre de la DCE, classe les types de cours d'eau selon leur contexte géoclimatique d'appartenance (HERs) et selon des critères de taille (depuis les plus petits cours d'eau permanents jusqu'aux TGCE). Cette classification de taille s'est essentiellement calée sur l'ordre de Strahler, qui est une vision indirecte de la taille de l'hydrosystème donnée par le niveau cumulé d'ordres de confluences depuis la source du drain principal jusqu'au site d'observation.

Si elle permet de bien classer les hydrosystèmes dans la région d'appartenance, selon des conditions naturelles qui y siègent, il a aussi été remarqué qu'à l'échelle inter-régionale ou inter-bassins de notre territoire français, **cette typologie peut donner une classification assez bruitée** par rapport à d'autres critères tout aussi légitimes que l'ordre de Strahler pour rendre compte d'une certaine hiérarchie de taille des hydrosystèmes (ex : longueur du cours d'eau depuis sa source, surface intégrée de bassin versant au site d'observation, etc...).

En effet, le coefficient de drainage d'un bassin versant au km² est assez variable en fonction du milieu naturel, des pentes, de la nature et de la perméabilité des substrats géologiques, du régime de pluies dans la région naturelle considérée etc.... Or le coefficient de drainage du bassin versant et les caractéristiques morphologiques du réseau des linéaires de surface vont influencer directement sur la relation taille-ordre.

D'autre part, les ordres de Strahler ont fait l'objet d'un calage initial pouvant présenter des différences selon les grands bassins hydrographiques. En effet, chaque Agence de Bassin a eu à construire, à un moment donné, un SIG de représentation et de hiérarchisation des hydrosystèmes de son propre bassin hydrographique (BD Carthage, dérivée de la BD Carto de l'IGN).

Or, les contextes naturels et les façons de procéder au sein de chaque Agence n'étant pas forcément identiques, chacune d'entre elle a pu définir les ordres des cours d'eau sur son territoire en prenant en compte de façon éventuellement décalée le début des très petits cours d'eau permanents. Ainsi, entre autres hétérogénéités inter-bassins plus ou moins marquantes, le Bassin Loire-Bretagne est celui où un ordre donné de cours d'eau est atteint pour des cours d'eau sensiblement plus petits que sur tous les autres grands bassins hydrographiques de France. D'autres hétérogénéités de typologie physique, bien que moins marquantes, ont aussi pu être relevées entre les autres bassins, au moins dans certains contextes naturels spécifiques (par exemple, cours d'eau du littoral Languedocien).

Dans ce contexte de typologie abiotique nationale non-homogène du point de vue du critère de taille, une sollicitation de participation à un groupe européen d'intercalibration diatomique des Très Grands Cours d'Eau qui, en fonction de l'état encore embryonnaire de la réflexion nationale à ce sujet, apparaissait comme une opportunité à saisir, est intervenue de façon anticipée par rapport au contexte de programmation du travail de fond sur ce sujet, soit au tout début du 2^{ème} semestre 2015. En effet, dans un contexte de bancarisation nationale et/ou de validation de données en retard sur plusieurs bassins, le chantier d'analyse de données était prévue pour commencer de 6 mois à 1 an plus tard.

La décision de participation nationale à cet exercice démarrant en période estivale (premier E-mail de sollicitation daté du 15 Juin 2015, avec une deadline de fourniture des éléments nécessaires en Septembre 2015), dans un contexte organisationnel non planifié au départ, a nécessité une bonne dose d'adaptation et une forte réactivité afin de pouvoir fournir sous un délai rapide :

- **1)** un jeu de données national d'évaluation TGCE comportant un assortiment représentatif de résultats d'évaluation diatomique au relevé, ainsi qu'un éventail fourni de données abiotiques d'accompagnement formatées très précisément selon la demande UE (donc pas forcément en bonne correspondance avec la pratique nationale habituelle des réseaux de surveillance) ;

- **2)** des réponses étayées à diverses questions sur la légitimité et l'efficacité de la méthode indicielle nationale (l'IBD₂₀₀₇) ;
- **3)** un premier prototype de dispositif d'évaluation calé avec encore très peu de recul et faisant appel à une dose importante d'empirisme.

En dépit de cette forte tension sur les échéanciers, la participation à cet exercice était très intéressante pour la France car elle présentait le double avantage :

- 1) de se placer directement dans le cadre d'un système d'évaluation basé sur un **niveau de référence par défaut déjà testé et approuvé par un consortium d'Etats-Membres et par l'UE** (en fonction des résultats obtenus dans le cadre d'un exercice d'intercalibration antérieur, coordonné les Allemands, qui avait été réalisé 3 ans plus tôt dans le cadre des travaux du X-GIG « Large Rivers »), évitant ainsi d'avoir à mener une réflexion spécifique et éventuellement décalée sur le sujet,
- 2) de pouvoir confronter d'emblée un **premier prototype de système national d'évaluation**, au calage très certainement perfectible, en comparaison directe avec ceux d'autres pays participants, puis de procéder ensuite aux **ajustements nécessaires** via plusieurs itérations, permettant ainsi de vérifier et de faire agréer la conformité de niveau d'évaluation au sein de l'exercice collectif.

Il a donc été décidé, en accord avec la Direction de l'Eau et l'ONEMA (devenu depuis AFB) de faire preuve de la réactivité maximale pour que la France puisse participer pleinement à cet exercice et bénéficier de l'expérience antérieurement capitalisée par d'autres Pays-Membres sur ces systèmes de grande taille démunis de situations de référence vraies.

2.2. Aspects particuliers, difficultés à résoudre :

Sur le plan des moyens humains et matériels, pour qu'une inscription effective de la France à cet exercice collectif soit possible dans des conditions très peu anticipées (*i.e.* sous moins de 3 mois, au titre d'une action non planifiée dans ces termes dans la fiche-action 2013-2015 en cours), il a fallu déployer une réactivité tout-à-fait particulière, qui a nécessité une importante ré-affectation de la force de travail de l'équipe Irstea/CARMA pour une re-priorisation rapide sur ce dossier, en avance d'un semestre sur le début effectif de la nouvelle fiche-action 2016-2018 et en anticipation de près d'un an par rapport à l'échéancier de travail effectivement programmé sur le plan de la réalisation de l'étude nationale TGCE. Cette action s'est donc forcément inscrite en compétition avec les échéanciers de rapportage d'autres actions en cours (notamment Fiche-action AQUAREF B « Incertitude IBD₂₀₀₇ »), ou prévues depuis le tout-début de la nouvelle fiche-action (notamment le travail sur de nouvelles métriques diatomiques de diagnostic), qui faisaient recours au même pool de ressources humaines permanentes et temporaires programmées et dont la réalisation et l'aboutissement ont subi et/ou subissent encore des retards.

Outre cette difficulté d'ordre programmatique, il est important de rapporter ici certaines difficultés spécifiques rencontrées du fait de la **différence de pratique** entre notre niveau national et certaines définitions typologiques qu'il nous était demandé de respecter strictement dans le cadre de l'exercice d'intercalibration européenne « Large Rivers », à savoir :

- la **typologie hydrochimique des cours d'eau**. Le document-guide Ecostat faisant référence pour l'exercice imposait de classifier la minéralisation des hydrosystèmes à partir de seuils d'alcalinité exprimé en milli-équivalents de carbonates /L d'eau sur moyenne pluri-annuelle, descripteur non-utilisé en France et dont le respect a nécessité un important post-traitement des données physico-chimiques.
- La **taille minimale des très grands cours d'eau** selon l'U.E., dont plusieurs sont pluri-nationaux à l'échelle de l'Europe (ex : Rhin, Danube etc...), se base sur une **surface cumulée de bassin versant supérieure ou égale à 10 000 km²** à la station d'observation, alors que notre typologie nationale utilise à la fois leur appartenance à une HER et un critère de taille différent, défini à partir des ordres de Strahler). Le respect de la typologie utilisée dans le cadre du GIG « Large Rivers », qui s'imposait à nous, a donc nécessité la réalisation à notre niveau d'un travail spécifique imprévu mobilisant une certaine dose d'expertise.

L'inscription de la France dans cet exercice collectif est intervenue par force et de façon non anticipée dans un créneau temporel posant des problèmes spécifiques de réactivité pour la plupart des personnes potentiellement impliquées de l'équipe de recherche CARMA (travail estival réalisé entre le 15 Juin et Septembre 2015, dans un contexte posant des problèmes de surcharge à la fois vis-à-vis de campagnes estivales d'échantillonnage hydrobiologique programmées assez longtemps à l'avance, et de planification de congés estivaux des agents tenant compte des impératifs de réalisation contractuelle à cette saison.

Du fait de la nécessité d'adopter une pratique homogène avec les autres pays-Membres, l'inscription dans ce travail a imposé la réalisation dans l'urgence d'un travail spécifique de requête, d'expertise et de post-traitement, à partir de l'assise de données immédiatement ou assez rapidement disponible, sur les 2 plans précités (caractérisation hydrochimique, critère de surface intégrée de BV utilisé pour le seuillage TGCE).

Le travail sur la typologie de taille a permis de soulever un problème de divergence assez importante entre les résultats de la typologie d'hydrosystèmes selon les critères UE et selon ceux de notre propre typologie nationale (ces critères représentant 2 manières *a priori* légitimes pour approcher la taille d'hydrosystèmes et pour les classer selon ce type d'information).

En effet, certains cours d'eau Français arrivent de reliefs proches et traversant le domaine littoral méditerranéen représentent probablement une exception typologique sur le plan du rapport entre surface de leur BV et leurs caractéristiques morphologiques (cours d'eau sous influence de phénomènes orageux de type Méditerranéen ou Cévenol). Entre autres exemples un peu moins marquants, ces particularités des environnements naturels et climatiques de cette zone du bassin hydrographique RMC ont par exemple conduit à classer l'Aude en TGCE dès l'aval de Carcassonne alors que sa surface intégrée de bassin versant à cette station (2 768 km²) représente à peine plus du quart du seuil inférieur de taille adopté par l'UE pour définir les TGCE. Inversement, certains cours d'eau du grand Bassin Parisien (zone à dominance de plateaux ou de plaines plutôt calcaires) dépassent déjà assez largement le seuil de surface de 10 000 km² utilisé par l'UE au moment où ils basculent dans le type TGCE de notre typologie nationale, dont plusieurs gros affluents de la Seine et la Seine elle-même (entre 12 000 et 13 000 km²) ; l'exemple le plus extrême de notre système national étant la Maine à Bouchemaine, encore qualifiée de GCE alors que sa surface cumulée de BV dépasse les 23 000 km² à cette station (soit 2,3 fois le seuil utilisé par l'UE).

Il est logique que ces 2 classifications typologiques, qui n'utilisent pas le même critère de décision, puissent ne pas donner une classification complètement cohérente. Cependant, il était obligatoire d'utiliser la typologie européenne pour s'intercalibrer et le résultat de l'intercalibration européenne doit être respecté en retour au niveau national pour appliquer de façon réglementaire un système d'évaluation DCE-compatible, ce qui rend préférable une harmonisation permettant de s'aligner pour une gestion plus homogène de ces cas extrêmes.

Sur le plan pratique, hors site particulier faisant l'objet d'un suivi de station hydrométrique, les données générales de caractérisation du site d'observation disponibles dans notre système de bancarisation nationale Sandre-SIE incluent rarement certaines données issues d'un traitement géomatique, en particulier les données de surface intégrée de bassin versant au site d'observation d'un réseau de surveillance, les données de linéaire cumulé du cours d'eau depuis sa source, les données de pentes de tronçons, de pente moyenne d'un cours d'eau depuis sa source etc... Or la typologie des cours d'eau européens élaborée par ECOSTAT, notamment pour cadrer les exercices d'intercalibration européenne des méthodes d'évaluation de l'état écologique, fait un recours systématique à un critère de surface intégrée de bassin versant au site d'observation, en croisement avec d'autres critères hydrochimiques ou de localisation géographique. Dans le contexte de réalisation de ce chantier, vu qu'un travail systématique de modélisation sous SIG, non planifié et impliquant de surcroît la participation d'une autre équipe, était impossible à mobiliser sous les délais requis, il a fallu mobiliser des données pré-existantes, disponibles seulement de façon approchée et incomplète via les référentiels DAHU et SYRAH, et adjoindre une part additionnelle d'expertise afin de pouvoir fournir les caractérisations de cours d'eau demandées au site d'observation.

Dans notre contexte national, il est aussi apparu utile d'étudier aussi l'éventualité d'un seuil de surface un peu abaissé par rapport au seuil d'intercalibration européenne de 10 000 km² (**un seuil > 8000 km² pouvait *a priori* apparaître comme un compromis national intéressant**) pour plusieurs raisons :

- La plupart des grands affluents des 5 fleuves Français avant leur confluence avec le drain principal sont déjà des hydrosystèmes de taille conséquente, pour lesquels le système d'évaluation diatomique actuellement en vigueur en France a été initialement calé de façon approximative du fait de la lacune nationale en situations de référence représentatives d'aussi grands cours d'eau. A ce niveau de taille, on tombe donc déjà dans la même problématique que celle qui a conduit à retarder la mise au point d'un système national d'évaluation TGCE. En effet, sur des hydrosystèmes de cette taille, il existe déjà une pression anthropique omniprésente d'une part, qu'elle soit d'origine agricole ou liée à l'implantation répétée de collectivités urbaines au long du drain considéré, et une absence de références vraies valides pour cette échelle de taille d'autre part. Or, sans forcément atteindre le seuil d'intercalibration européenne de 10 000 km², qui a probablement été raisonné en partie en fonction de l'enjeu particulier attaché à l'homogénéité d'évaluation de cours d'eau trans-frontaliers à l'échelle de l'Europe et qui vise surtout à garantir un certain niveau de comparabilité des hydrosystèmes soumis au même exercice d'intercalibration, la plupart des affluents principaux de nos 5 grands fleuves font plus de 8 000 km² et ont au moins une station de surveillance implantée au-dessus de ce seuil de taille avant leur confluence avec le drain principal.

- Il en est de même pour certains fleuves côtiers déjà de grande taille, par exemple la Charente, la Vilaine et l'Adour qui se jettent directement en mer, ainsi que pour certains grands cours d'eau à enjeu transfrontalier qui quittent ensuite la France (Meuse, Moselle), mais qui n'atteignent pas tout-à-fait 10 000 km² à la station de réseau la plus aval de notre territoire national.

2.3. Etat des données nationales TGCE / diatomées :

2.3.1. Historique et portée des requêtes :

Afin de procéder à l'inventaire et au rapatriement des doublets de données (hydrochimiques + diatomiques) disponibles et utilisables en France pour avancer sur cette problématique TGCE, le principe général adopté a été de permettre l'exploitation de requêtes seules de la base nationale Pandore visant 2 applications différentes :

- sites de cours d'eau français situés sur des bassins versants intégrés de **plus de 10 000 km²** (autorisant ainsi l'utilisation des données de ces sites dans le cadre de l'exercice d'intercalibration européen « Large Rivers »)
- sites de cours d'eau français situés sur des bassins versants intégrés de **plus de 8 000 km²**, afin d'obtenir l'assise maximale de données permettant de retravailler et consolider notre dispositif national d'évaluation des TGCE, et d'étudier l'opportunité éventuelle de prendre en compte ce seuil.

En pratique, ces requêtes ont dû intervenir en 2 temps, en fonction d'éléments contextuels relatifs aux impératifs temporels sensiblement différents de programmation et de restitution des actions :

- Une première échéance temporelle assez courte s'imposait à nous, à savoir la fourniture dans les meilleurs délais du jeu de données national TGCE soumis à l'exercice d'inter-calibration européenne (deadline de fourniture : Septembre 2015). **Une première série de requêtes a été réalisée entre fin Juin et fin Août 2015 dans PANDORE**, basée uniquement sur les données déjà disponibles à l'époque et autorisant un couplage entre données abiotiques d'accompagnement et relevés diatomiques. Elle a permis de satisfaire aux demandes de court terme du GIG « Large Rivers », et en particulier de lui fournir dans les temps requis le jeu de données Français, soumis au processus d'intercalibration en date du 15 Septembre 2015 (pour une deadline imposée à fin Septembre 2015). Les données ainsi récupérées ont aussi permis de produire le 1^{er} rapport d'avancement d'étape de cette action « Evaluation des TGCE », restitué à l'ONEMA en Janvier 2016, par anticipation par rapport au contenu de la fiche-action 2016-2018.
- La réalisation ultérieure d'une 2^{ème} requête plus complète, permettant une prise de recul plus importante par rapport aux référentiels nationaux disponibles et l'actualisation de certains résultats de l'étude nationale « Evaluation des TGCE ». En fonction de la complétion de la BDD PANDORE, réalisée au fil de l'eau par nos collègues Irstea de Lyon, et de la date d'arrivée d'un nouveau CDD en biostatistiques dans notre équipe (David CARAYON, début de contrat en date du 01-07-2017), **cette requête la plus actualisée possible a été réalisée en Juillet 2017** en vue de sous-tendre les prises de décisions nationales relatives à la nouvelle version d'Arrêté Evaluation en préparation à l'époque (signée du 27 Juillet 2018 et publié au JORF en date du 31-08-2018) et d'actualiser les résultats d'étude nationale TGCE fournis dans le présent rapport final.

2.3.2. Données de description générale des sites (stations) :

Sous Pandore, les données attachées aux stations de prélèvement et/ou de mesure comprennent diverses informations générales caractéristiques, dont le nom du cours d'eau, le bassin hydrographique, une numérotation nationale unique de station de réseau (numéro Sandre), le nom du cours d'eau concerné (ex : la Dordogne), la description résumée de l'implantation du site (ex : pont de la D84 à Cénac), ses informations géoréférencées, son altitude, sa surface intégrée au point de mesure etc...

Comme déjà mentionné auparavant, la participation de la France à l'exercice Européen d'intercalibration « Large Rivers » nous imposait de fournir une liste spécifiée et précise de données relatives à nos TGCE Français et aux sites de réseaux les concernant, dans le cadre de tableaux standardisés devant être remplis de façon complète et homogène par tous les pays souhaitant participer à l'exercice.

Or malheureusement, certaines caractéristiques demandées, relatives aux cours d'eau et aux sites de réseaux qui les concernent, ne sont pas systématiquement disponibles ou accessibles dans l'environnement national Sandre – SIE. Fait relativement incompréhensible à notre époque dans le cadre d'un référentiel national « Données », de tels manques sont encore constatés alors qu'ils concernent des attributs de sites présentant une certaine stabilité dans le temps, voire déterminables une fois pour toutes. C'est notamment le cas en matière d'attributs de sites qui nécessitent un traitement un peu spécifique d'informations géographiques faisant éventuellement recours à des algorithmes de calcul (par exemple pentes moyennes de cours d'eau depuis la source) ou de modélisation d'interfluves (par exemple calcul de surface intégrée de bassin versant à un point d'observation donné).

La fourniture des données demandées étant obligatoire pour que le GIG puisse s'assurer du respect de la typologie d'hydrosystèmes utilisée dans le cadre de l'exercice, il a été nécessaire de mobiliser de l'information existante, éventuellement incomplète, imprécise ou perfectible, afin de répondre aux exigences de l'exercice. Ainsi, 2 sources d'information relatives à des surfaces intégrées de bassins versants, acquises ou modélisées en fonction d'objectifs différents, étaient disponibles dans 2 champs différents de la BDD PANDORE, ces informations non exactement équivalentes présentant assez souvent des défauts de cohérence relative, ainsi qu'une lacune pour certaines stations situées très à l'aval des bassins versants, souvent implantées sur très grands cours d'eau et/ou réseaux de canaux côtiers :

- des données assez anciennes (dites surfaces DAHU), dont la genèse date des années 2003-2007 (origine Cemagref Lyon / LHQ), qui ont été utilisées à l'époque pour l'élaboration de la typologie nationale des cours d'eau et dans le cadre des premiers exercices d'intercalibration européenne sur les compartiments MIB, poissons, diatomées. La complétion de ces référentiels était plutôt exhaustive sur les cours d'eau de taille petite, moyenne et grande, mais était de plus en plus lacunaire au fur et à mesure que les cours d'eau se rapprochent de leur exutoire, les zones de très faibles pentes présentant des difficultés particulières de délimitation des tracés d'interfluves. Cette difficulté d'ordre général était encore accrue du fait de la faible résolution verticale du MNT disponible à l'époque. Cet état de fait n'a donc pas vraiment posé de problème dans le cadre des premiers exercices d'intercalibration européenne, qui visaient des hydrosystèmes de plus petite taille. Il en pose sensiblement plus sur les TGCE qui sont renseignés de façon de plus en plus lacunaire lorsqu'ils s'écoulent dans des zones à très faible pente, ce qui est souvent le cas lorsqu'ils arrivent à proximité de la zone littorale et de leur exutoire.
- des données modélisées plus récemment dans le cadre des approches hydromorphologiques SYRAH (dites surfaces SYRAH) qui, outre un reliquat de difficultés rencontrées en liaison avec les aspects de résolution altitudinale (avec une version de MNT cependant plus précise qu'à l'époque antérieure), étaient porteuses d'une difficulté d'une autre nature car elles n'indiquent pas une surface cumulée de BV au site de réseau, mais à l'aval de la masse d'eau à laquelle appartient le site.

A l'époque de cette sollicitation et dans le réseau de contraintes du moment, il n'était pas possible à nos collègues du pôle Irstea-ONEMA de Lyon, compétents dans ce domaine géomatique et disposant des outils spécifiques nécessaires (SIG, MNT, techniques de modélisation géomatique etc...), de réviser ou de définir des surfaces de BV de tous les sites à information incomplète ou présentant une incohérence apparente entre les 2 sources d'information disponibles (tâche conséquente, non prévue dans leur plan de charge contractuel). Dans le but de nous aider à partir d'informations pré-existantes, ils ont donc procédé à une requête complémentaire permettant à moindre coût de compléter certains champs vides, en retrouvant quelques analyses de surfaces intégrées dont ils avaient pu avoir besoin par le passé. Mais, dans les délais tendus imposés pour la participation nationale à cet exercice d'intercalibration, il était exclu de les faire procéder à de nouvelles délimitations des surfaces par modélisation rigoureuse sous SIG.

Nous avons donc eu recours à quelques requêtes complémentaires au cas par cas (via le nom du cours d'eau, l'altitude de la station, ses coordonnées géographiques) pour repérer toutes les stations de réseaux disponibles dans le domaine des Très Grands Cours d'Eau et ayant fait l'objet d'une codification Sandre, ainsi que leur agencement respectif amont-aval. Un post-traitement de ces données a ensuite été réalisé, sur la base des informations de ces 2 référentiels DAHU et SYRAH et, chaque fois que possible, sur la base d'informations complémentaires provenant d'autres supports disponibles (fiches-stations SANDRE, informations collectées sur des sites hydrométriques utilisés en modélisation quantitative des débits, etc...).

Ces données annexes, requêtées de façon ciblées et au cas par cas pour combler des manques repérés, ont notamment permis de stabiliser les informations de surfaces cumulées, d'incorporer les données de débit modulaire annuel (information également demandée dans le cadre de l'exercice) et de compléter certains autres manques (ex : N° de masse d'eau...), ainsi que de corriger certaines erreurs constatées dans les données existantes (par exemple, les surfaces de certains grands cours d'eau inter-frontaliers donnés dans la base, comme le Rhin et le Rhône, n'étaient pas cumulés depuis la source mais depuis l'entrée du cours d'eau sur le territoire Français).

L'agencement amont-aval des sites de suivi des TGCE a été vérifié et est parfaitement valide. Les surfaces finalement stabilisées puis utilisées dans ce rapport et dans les jeux de données nationaux envoyés au GIG « Large Rivers » incluent une part d'expertise et un degré d'imprécision modéré, mais variable selon l'origine réelle de la donnée et son mode d'obtention originel. En effet, une petite proportion d'entre elles a été donnée par estimation approchée, en veillant à bien respecter l'agencement des surfaces en fonction de la disposition respective amont-aval des sites situés sur un même linéaire de cours d'eau, et en tenant compte de l'emplacement précis des confluences avec des affluents apportant une contribution significative de surface cumulée.

2.3.3. Données biologiques (données diatomiques) :

Cette étape du travail a visé à repérer, sur la base de la définition des sites TGCE faite juste auparavant, la disponibilité de relevés diatomiques jugés valides sur les stations éligibles au statut TGCE. C'est sur cette base des relevés diatomiques existants et de leur date de réalisation que des requêtes et travaux de post-traitement de données abiotiques de qualité d'eau (3^{ème} étape de cette sélection) présentaient leur utilité.

Par rapport à sa formulation initiale datant de 1996 (normalisée AFNOR 2000), l'IBD a subi une évolution interne notable en vue de la révision de la norme AFNOR NF T 90-354, qui est intervenue début 2007. Dans l'utilisation plus ancienne de l'IBD normalisé (version dite « IBD₂₀₀₀ »), cet indice faisait recours à seulement 209 profils d'espèces constitutives, dont 173 espèces vraies et 36 espèces appariées. A partir de 2007, l'indice a fait appel à 834 profils de taxons vrais et à des profils de taxons altérés pour chaque frustule dont l'observation révélait des déformations tératologiques, le but étant d'altérer la note en cas de déformations (suspicion de pollution toxique lorsque le dénombrement de ces formes dépasse le faible pourcentage pouvant être observé en conditions naturelles). Même si les 2 versions successives d'IBD ne conduisent pas à des différences importantes d'évaluation de qualité, il s'ensuit une certaine rupture de continuité et de comparabilité, en premier lieu dans l'identification des inventaires diatomiques (le guide d'identification attaché à la méthode ayant changé à cette même époque...), puis dans le calcul de l'indice (assise nettement accrue de taxons contributifs, raffinement d'affectation des profils de qualité existants).

Pour l'élaboration du nouveau dispositif d'évaluation TGCE, il a donc été décidé de ne prendre en compte que les inventaires diatomiques collectés et identifiés à partir de 2007 dans les différents réseaux nationaux accessibles (RCS, RREF, RCO, RCB).

Outre les critères de taille minimale du bassin versant au site (la requête ayant conduit à identifier tous les relevés diatomiques collectés sur **154 sites de plus de 8 000 km²** (dont **99 sites de plus de 10 000 km²** susceptibles d'être retenus dans le jeu France pour l'exercice européen d'intercalibration « Large Rivers »), il a été procédé à une sélection plus restrictive éliminant les relevés non prélevés entre Avril et Septembre (relevés s'éloignant trop des conditions de prélèvement normatif de belle saison prônés dans le cadre des réseaux de surveillance classiques).

De même, les relevés s'éloignant trop de la pression d'observation-cible recommandée dans le protocole normatif (400 valves +/- 20 %) ont été éliminés, certains d'entre eux, à dénombrements nettement plus nombreux (jusqu'à 800 et même 1200 valves dénombrées), émanant probablement d'études conduites avec un protocole spécifique, et d'autres présentant un défaut de fiabilité en fonction d'une assise de dénombrement sensiblement insuffisante.

2 requêtes ont successivement été réalisées selon ce cadrage général :

- La première, réalisée en début d'été 2015 dans l'optique d'une fourniture rapide du jeu de données d'intercalibration. Cette première requête a conduit à repérer 645 relevés diatomiques *a priori* utilisables (sous réserve de l'existence ultérieure de données de suivi chimique convenables), collectées sur 154 sites différents de plus de 8 000 km². Cette requête permettait à la fois de produire de la donnée compatible avec le seuil Européen de 10 000 km² (il en a été extrait notre jeu de donnée national d'intercalibration, constitué de 210 relevés diatomiques représentatifs de 99 sites de réseau de plus de 10 000 km²), et de procéder à l'étude nationale TGCE envisageant et inter-comparant les 2 seuils de taille-candidats.
- La seconde, réalisée en Juillet 2017 à l'ultime date possible par rapport au bon achèvement de cette étude nationale. Cette seconde requête incluait les plus récents inventaires diatomiques dernièrement incorporés dans la base de données PANDORE (complétion totale des années 2012 et 2013 au niveau national) et pouvant d'emblée être correctement couplés à des données abiotiques de chimie des eaux. L'effectif concerné était de 667 relevés couplés utilisables, acquis sur 127 sites différents de réseaux de surveillance cumulant plus de 8000 km² de bassin versant intégré au site d'observation. Comme la précédente, cette requête permettait de traiter et d'inter-comparer l'évaluation des TGCE selon les 2 hypothèses de seuils de taille - candidats.

2.3.4. Croisement avec les données de qualité d'eau, bilan quantitatif :

La première requête croisée sur la BDD PANDORE, qui visait à associer les relevés diatomiques TGCE d'emblée disponibles aux OPECONT-Chimie correspondantes, a été lancée en début de **2^{ème} semestre 2015**, dans le premier objectif de fournir dans les meilleurs délais la sélection d'un jeu de données national représentatif répondant aux critères typologiques requis (en particulier, le seuillage de 10 000 km²) et servant de base à l'exercice d'intercalibration européenne, qui constituait l'urgence absolue du moment. En effet, la fourniture de ce jeu de données en Septembre 2015, dans le respect d'un certain équilibre quantitatif vis-à-vis des données d'autres pays-membres, était une condition *sine qua non* de notre participation nationale.

D'autre part, cette première requête a été pensée pour couvrir aussi les besoins en données nécessaires au démarrage de réalisation de l'étude nationale TGCE.

Cette première requête exercée dans ce contexte temporel très contraint ont permis de dresser le constat que la base de données était renseignée de façon plutôt satisfaisante sur le plan des inventaires diatomiques (données biologiques incluant déjà la quasi-totalité des relevés 2013 à l'échelle française), mais que l'inclusion des données correspondantes de chimie des eaux de plusieurs bassins était plus en retard, avec les données de certains bassins complètement manquantes sur toute l'année 2013, voire même sur 2012 et 2013, pouvant faire perdre le bénéfice de **2 années de suivi diatomique** sur ces bassins.

La **Figure 1** ci-dessous illustre le problème rencontré à l'époque pour le couplage entre les relevés diatomiques (période de référence retenue : de 2007 à fin 2013) et des données de chimie disponibles, selon la méthodologie proposée par le GIG « Large Rivers » (mise en relation données abiotiques-relevés diatomiques sur la base d'une **intégration annuelle** des données de chimie). Rappelons que l'assise de données utilisée pour asseoir la démarche d'évaluation TGCE s'est appuyée sur les inventaires diatomiques collectés dans ces cours d'eau à partir de l'année 2007, dans l'application de la norme d'acquisition NF T 90-354 révisée 2007, afin de garantir une meilleure homogénéité des pratiques de terrain, du cadrage d'acquisition taxonomique et par conséquent, une comparabilité plus satisfaisante du référentiel utilisé.

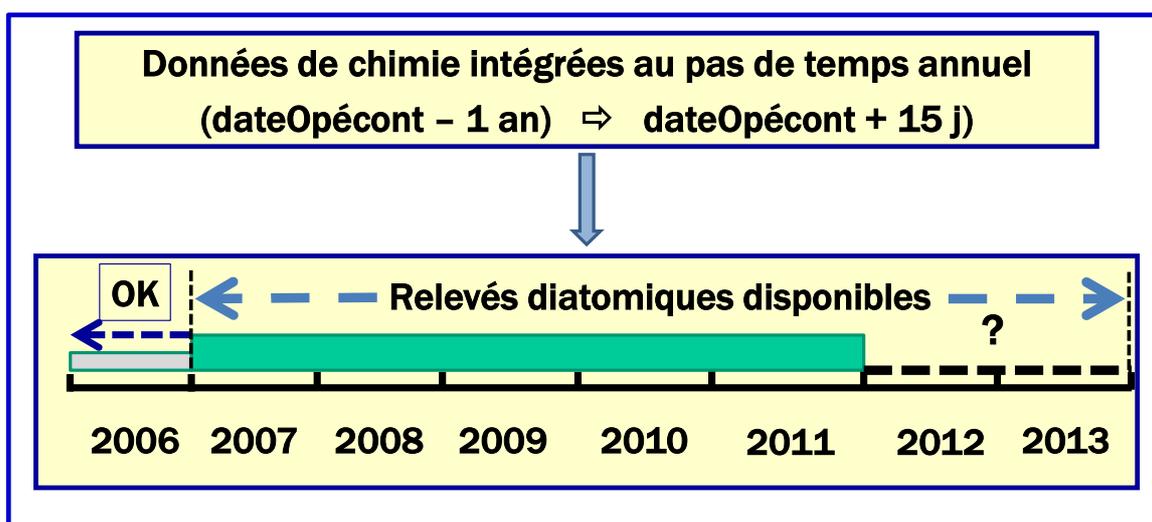


Figure 1 : Problèmes rencontrés pour la mise en relation entre relevés diatomiques TGCE et données abiotiques de chimie (requêtes PANDORE 2015)

Dans le contexte où le nombre de sites de réseaux sur TGCE est relativement limité par rapport à l'assise capitalisée sur cours d'eau plus petits, le décalage de disponibilité constaté à l'époque entre les inventaires diatomiques (jusqu'à 2013 inclus) et les données hydrochimiques correspondantes (non encore disponibles après 2011 pour plusieurs bassins hydrographiques) conduisait à faire perdre le bénéfice de plus de 140 inventaires diatomiques capitalisés lors des 2 dernières années du créneau temporel visé, représentant près de 20% des relevés TGCE disponibles sur une période de référence 2007-2013.

Le problème se posait d'ailleurs exactement dans les mêmes termes pour le couplage avec une chimie intégrée **au trimestre**, pratique *a priori* plus convenable pour résumer le pas de temps caractéristique d'intégration biologique réalisé à l'échelle d'un prélèvement de biofilm diatomique (de l'ordre de 2 à 3 mois) ; ce pas de temps plus réaliste étant par ailleurs retenu pour la réalisation finale de l'étude nationale TGCE.

Ces données complémentaires de chimie, de même que quelques prélèvements diatomiques manquants, ayant fait l'objet d'une complétion progressive au fil de l'eau, il a été procédé, en début du 2^{ème} semestre 2017, à une **nouvelle requête de la base PANDORE** actualisée sur le pas de temps **2007-2013**, afin de pouvoir disposer de tous les doublets de données complets mobilisables à cette date et suffisamment comparables sur le plan de la norme d'acquisition (voir **Figure 2** en page suivante).

En pratique, cette nouvelle requête arrivant à la date la plus tardive possible pour pouvoir achever l'étude nationale « Evaluation TGCE » ne pouvait plus être mise au service de l'exercice européen d'intercalibration, dont le rapport final avait déjà été envoyé à l'U.E. un an auparavant (soit en date du 21-06-2016). Par contre, cette assise de données augmentée a été utilisée pour ré-actualiser les résultats obtenus dans le cadre de l'étude nationale, en particulier les relations pression-impact, ainsi que la simulation des résultats d'évaluation obtenus sur l'ensemble des sites TGCE selon 4 scénarios-tests associant, selon des modalités un peu différentes, des possibilités d'options typologiques TGCE et l'application des grilles d'évaluation intercalibrées).

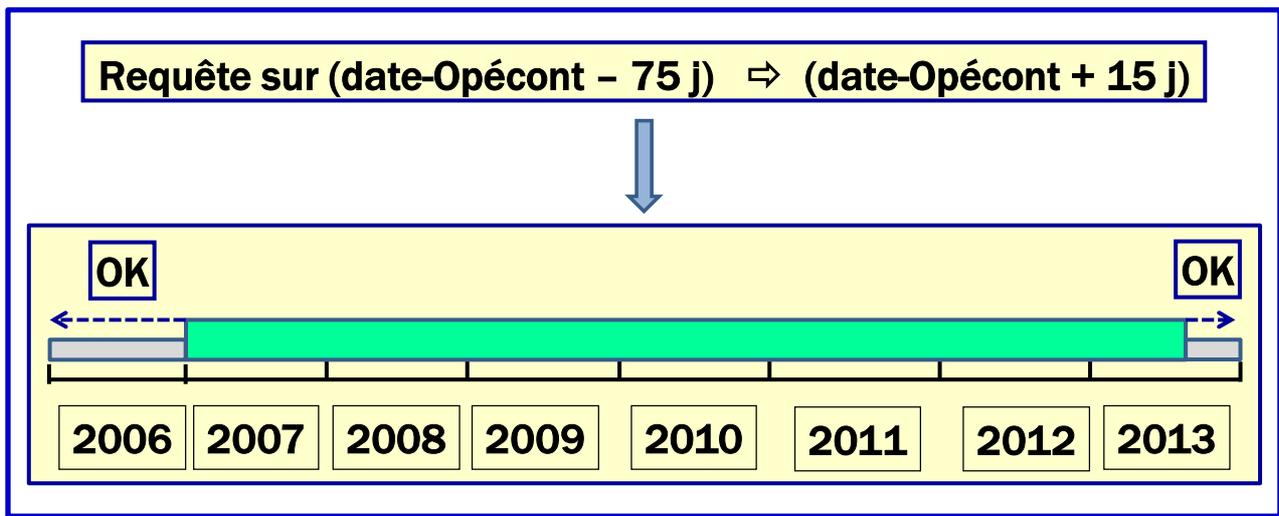


Figure 2 : Mise en relation entre relevés diatomiques et données abiotiques de chimie pour l'étude nationale TGCE (requêtes PANDORE 2017)

Le **Tableau 1** ci-dessous récapitule les aspects quantitatifs globaux relatifs aux données disponibles pour les différents aspects de l'étude TGCE, en fonction de la date des 2 requêtes successives PANDORE.

Tableau 1 : Description quantitative globale des données nationales disponibles selon la date de requête et le type d'utilisation (étude nationale versus Intercalibration Européenne)

Taille TGCE	Relevés diatomiques 2007 - 2013 disponibles (Sept. 2015)		dont couplage chimie (Sept. 2015)		Relevés avec couplage chimie (Juil. 2017)		Sélection Intercal EU (Sept. 2015)	
	nb. sites	nb. relevés	nb. sites	nb. relevés	nb. sites	nb. relevés	nb. sites	nb. relevés
Plus de 10 000 km ²	99	532	97	433 *	99	550	95	221
Plus de 8 000 km ²	154	645	122	523 *	127	667	NA	NA
Différentiel	55	113	25	90	28	117	NA	NA

* : quasi-absence de données de chimie 2012 et 2013 requêtables dans Pandore

Les **colonnes à fond rose** en partie gauche du tableau représentent le résultat du **premier train de requêtes PANDORE** faites à marche forcée au début du 2^{ème} semestre 2015, en phase avec les échéanciers de l'exercice d'intercalibration européenne.

Pour les relevés TGCE acquis sur des bassins versants de **plus de 10 000 km²**, par rapport aux relevés diatomiques disponibles, on peut y constater le **déficit de (532 – 433) = 99 relevés** aux données de chimie manquantes.

Pour les relevés acquis sur des bassins versants de **plus de 8 000 km²**, ce déficit est porté à $(645 - 523) = 122$ **relevés**. Dans les 2 cas, ce sont 19 % des relevés diatomiques déjà capitalisés sur le pas de temps de référence qui ne peuvent pas encore être couplées en temps réel aux données de chimie correspondantes, appauvrissant assez substantiellement l'assise des relations qui peuvent être établies entre variables d'état relatives à la pression anthropique et résultats d'évaluation diatomique.

Par rapport à ce premier train de **requêtes 2015**, l'arrivée du reliquat de données nationales de chimie et de quelques données diatomiques en retard a permis de compléter le contenu de la partie abiotique de la base de données PANDORE sur les années 2012 et 2013. Du fait de cette complétion intervenant sur un jeu de données TGCE dont l'assise antérieure était plutôt modeste, il a été jugé intéressant, pour mieux asseoir la démarche nationale, de procéder à une **réactualisation de la requête PANDORE en Juillet 2017**, dont le récapitulatif est présenté dans les **colonnes à fond blanc** en début de partie droite du tableau.

Cette réactualisation a été d'emblée plus ciblée, puisqu'elle n'a sélectionné que des relevés diatomiques pouvant être associés avec au moins une OPECONT-chimie.

Pour les **TGCE de plus de 10 000 km²**, elle n'a pas fait évoluer le nombre de sites pris en compte (**99**), mais a légèrement augmenté le nombre de relevés diatomiques disponibles sur ces sites (550 au lieu de 532, soit + 3%). La principale marge de progression est venue de la complétion de données de chimie concernant des relevés diatomiques 2012 et 2013 déjà bancarisés ; les 2 sources cumulées d'augmentation **faisant progresser le total de relevés complets disponibles de 27 %** (soit 550/433).

Pour les **TGCE de plus de 8 000 km²**, par rapport aux 154 sites sur lesquels au moins un relevé diatomique était bancarisé, le nombre de sites pour lesquels il était possible de coupler des relevés diatomiques et des données hydrochimiques sur une base relativement solide s'est réduit à 127 (soit une perte de 18% de sites), mais des informations couplées de relevés diatomiques et de chimie concernent 5 nouveaux sites, et le **total de relevés complets disponibles a augmenté de 27,5%** (soit 667/523) **pour ce seuil de taille**.

Les colonnes en bleu figurant à droite du Tableau 1 contiennent le résumé des données transmises au GIG « Large Rivers » en 2015, dans le cadre du deuxième l'exercice d'intercalibration européenne (intitulé Fit-in exercise) auquel nous avons participé. Dans le cadre de cet exercice initié vers la mi-2015, la coordinatrice a demandé de procéder à une sélection permettant à la fois, parmi notre jeu de données de France, d'assurer une bonne représentativité nationale des gradients d'altération et du niveau d'évaluation, mais aussi de rester dans le respect d'un certain équilibre quantitatif de données par rapport à d'autres pays-membres les plus comparables (ex :Italie, Espagne) ; ceci dans le contexte où d'autres pays-membres participant à l'exercice pouvaient n'avoir qu'un seul TGCE transfrontalier sur leur territoire (exemple de la Meuse pour la Belgique / Flandres et du bassin du Danube, unique pour la Bulgarie et la Roumanie ou très dominant pour la Hongrie). Le nombre indicatif de données nationales que la coordinatrice recommandait de fournir pour la France devait être de l'ordre des 200.

Dans ce contexte, il a été procédé pour la mi-Septembre 2015 (échéance à respecter pour pouvoir participer à l'exercice), à l'élaboration d'une sélection de données Françaises visant à refléter de façon la plus représentative possible :

- ▶ notre assise de données nationales. A cet effet, la totalité des sites régulièrement suivis dans le cadre des réseaux a été représentée (soit 95 sites sur les 99 ayant bénéficié à un moment ou à un autre d'un suivi abiotique et diatomique, incluant depuis les sites des meilleures qualités jusqu'aux sites les plus altérés),
- ▶ l'amplitude maximum des situations à évaluer. En effet, s'agissant de la validation d'un système d'évaluation et y compris pour un site particulier, il était intéressant de représenter les gradients de situation abiotique et biologique les plus larges possibles.

Dans ce contexte et chaque fois que possible, à partir des doublets complets de données couplées (chimie-assemblages biologiques) disponibles à cette date pour chaque site suivi de façon à peu près régulière, il a été pris la décision d'envoyer la donnée d'évaluation biologique la plus basse, la plus haute et la donnée correspondant à la médiane pour ce site, ainsi que les données hydrochimiques en correspondance (moyenne annuelle calculée sur l'année antérieure à la date de prélèvement diatomique, conformément à la demande du GIG).

Le jeu de données représentant cette sélection nationale, construite selon les principes qui précèdent, a été envoyé une fois pour toutes vers la mi-Septembre 2015. Comme indiqué dans les colonnes à fond bleu en partie droite du Tableau 1, cet envoi comprenait 221 relevés diatomiques prélevés à 95 sites TGCE différents, ainsi que leurs données abiotiques et typologiques d'accompagnement.

2.4. : Post-traitement et complétion des données abiotiques

2.4.1. Données de qualité d'eau (aspects généraux)

Une tâche générique très importante dans cette étude consistait à assortir à un maximum de relevés diatomiques disponibles, de façon la plus représentative possible, les données de physico-chimie et d'hydrochimie correspondantes. Réalisée à 2 reprises, soit entre Juillet et Septembre 2015 pour les besoins de court terme liés à l'exercice d'intercalibration « Large Rivers », puis en Juillet 2017 pour l'actualisation de l'étude nationale TGCE, cette tâche visait les objectifs suivants :

- permettre une bonne description des données abiotiques en correspondance avec les inventaires diatomiques TGCE,
- être en mesure d'envoyer dans les meilleurs délais des jeux de données nationaux sélectionnés au GIG «LR», dont la description des paramètres d'accompagnement (du site ou du relevé selon le cas) soit conforme au contenu des tableaux (templates) de référence fournis.
- plus largement, sur le plan national, procéder à l'étude de relations paramètres abiotiques-flore, ou de relations pression-impact entre niveau de nutriments et réponse de notre système d'évaluation diatomique national TGCE (utilisant l'IBD₂₀₀₇ et des EQRs-IBD₂₀₀₇ par grand ensemble naturel homogène).

A ce stade de l'action, il a donc été nécessaire de s'assurer de la correspondance possible entre les relevés diatomiques TGCE extraits de la base Pandore et l'existence d'opérations de contrôle permettant de calculer des moyennes des conditions physico-chimiques et chimiques à ce site et pour ce relevé sur un pas de temps donné. Comme déjà indiqué, le GIG demandait des données abiotiques d'accompagnement des relevés diatomiques calculés en moyenne annuelle. Sur le plan de l'écologie de ce maillon-clé, cette façon de faire n'est probablement pas la plus convenable. En effet, sur le compartiment des diatomées benthiques, le relevé biologique est composé de l'échantillonnage intégré de biofilm prélevé sur plusieurs substrats durs. Cet édifice biologique ayant une durée de vie relativement limitée sur ces substrats, et sa composition spécifique est plus déterminée sous l'influence dominante des 2 à 3 mois de chimie saisonnière intégrée que d'une chimie moyenne annuelle.

Néanmoins, la demande du GIG portait sur des moyennes annuelles de chimie, demande probablement formatée pour atténuer la difficulté liée à des référentiels de données abiotiques assez lacunaires dans certains pays-membres de l'UE, ainsi que pour faciliter la participation des Pays-Membres aux exercices d'intercalibration, tous maillons confondus (il est en effet plus facile, que ce soit pour un aspect d'assise générale de données ou sur le plan de la réactivité nécessaire des acteurs concernés, de fournir des données moyennes annuelles calculées une fois pour toutes que des données rattachées à un créneau temporel plus étroit déterminé par la date précise d'un prélèvement biologique). Dans le cadre de cet exercice, nous avons donc respecté le principe de post-traitement des données demandé pour rester homogènes avec la pratique des autres pays participants.

Vu l'urgence temporelle de la demande des données nationales par le GIG et compte-tenu des limites rencontrées à l'époque vis-à-vis de la bancarisation des données PANDORE, nous nous sommes attachés à donner une réponse opérationnelle dans les meilleurs délais en requêtant rapidement les données de physico-chimie disponibles de 2007 à fin 2011 sur les sites où existaient des relevés diatomiques, le principe d'association d'informations suivi visant à relier les inventaires diatomiques et les Opécont-Chimie sur une base de moyenne annuelle, le créneau d'intégration temporelle précisément retenu étant indiqué en haut de la Figure 1.

Selon le principe mis en oeuvre, il a été tenu compte de toutes les Opécont-Chimie réalisées dans l'année précédant la date d'inventaire diatomique, ainsi que de celles intervenant moins de 2 semaines après. La raison de cet élargissement modéré de la fenêtre temporelle mise en pratique tient au fait que, dans des conditions hydrologiques stabilisées correspondant le plus souvent aux conditions de prélèvement diatomique en étiage estival, et vu que peu de relevés de chimie incluaient la longue liste d'analytes demandés par le GIG, un relevé de chimie complet intervenu moins de 15 jours après la réalisation du prélèvement diatomique a des chances de représenter tout aussi bien, voire mieux, les conditions *in situ* de l'épisode biologique concerné que des prélèvements plus éloignés dans le temps concernant l'année-même ou l'année précédente, réalisés de plus à des saisons diverses (de hautes eaux hivernales à des basses eaux estivales) ; d'autant que, pour certains réseaux, il peut n'être réalisé que 4 relevés de chimie pour l'année.

Une fois la fusion des données et un certain nombre de post-traitements réalisés sous R, nous avons réalisé un peu tardivement qu'il aurait convenu de requêter aussi les données de physico-chimie du 2^{ème} semestre 2006 pour assurer plus rigoureusement le regroupement chimie-biologie 2007 tel qu'indiqué. Cependant, le contexte du moment, beaucoup trop tendu pour retarder la fourniture des jeux de données nationaux participant à l'exercice d'intercalibration, n'a pas permis la rectification immédiate de cet aspect (*) dont la seule conséquence, tout-à-fait marginale, a concerné l'expression des relations nationales pression-impact fournies à titre indicatif à l'appui de la performance d'évaluation de l'IBD₂₀₀₇. De même, du fait de la complétion encore imparfaite, voire de la lacune en données abiotiques 2012 et 2013 sur certains bassins à cette époque, afin de simplifier l'approche, d'éviter d'avoir à réaliser des vérifications d'assise-données au cas par cas et par souci d'homogénéité globale inter-bassins dans ce contexte d'échéancier très court, les relevés nationaux fournis pour toute la France n'ont concerné que le créneau temporel [2007-2011].

() Nota : Il y a été remédié ultérieurement dans le cadre de la réalisation de la partie d'étude nationale visant à l'établissement de relations pression-impact basées sur une intégration annuelle des données physico-chimiques. Dans ce même cadre, il a aussi été utilisé une agrégation des données selon un principe trimestriel, plus convenable pour ce maillon biologique, selon le principe présenté en haut de Figure 2.*

La réalisation de cette tâche générale de couplage entre données biologiques et données abiotiques d'accompagnement a nécessité la résolution concrète de certains problèmes spécifiques, dont la nature et les solutions apportées sont décrites plus en détail dans les paragraphes suivants.

2.4.2. Gestion du problème d'absence de données numériques réelles

Dans les bases de données des Bassins et par contrecoup dans PANDORE, 3 colonnes se réfèrent au même paramètre analytique à une date donnée : une colonne contient une valeur du paramètre, une autre un code-remarque permettant d'interpréter le résultat de cette valeur. Enfin, une troisième colonne donne l'unité d'expression du paramètre. La valeur numérique contenue dans la première colonne évoquée, malgré qu'elle se place dans la même colonne, n'a donc pas toujours le même statut. Il peut s'agir soit d'une valeur quantifiée obtenue par analyse, soit du seuil de détection d'une méthode analytique (SD non atteint), soit du seuil de quantification d'une méthode analytique (SQ non atteint). Malheureusement, concernant l'archivage des données de bassin réalisé historiquement, il existe d'assez nombreuses cases vides dans les colonnes 2 et 3, qui propagent une incertitude sur la nature réelle de l'information figurant dans la première colonne.

Cette logique de construction, compréhensible à l'époque où le format des bases de données était très contraint et limitant, est défectueux et source de nombreux problèmes à l'époque actuelle, car il faut se référer au contenu de 3 colonnes différentes, pas toujours renseignées complètement selon les bassins, les époques, l'opérateur de Marché Public et éventuellement l'opération de rapatriement, pour savoir quel statut donner à la valeur numérique de la première d'entre elles ; dans un contexte où les seuils analytiques peuvent être assez largement variables selon l'époque de réalisation de l'analyse, les termes de référence du marché Public correspondant, le laboratoire, la méthode analytique... De plus, pour diverses raisons qui peuvent tenir à l'origine multi-sources, à des subtilités d'ordre typographique au moment des requêtes ou à des aspects de compatibilité inter-systèmes pour le langage de requête, quelques problèmes ou loupés, qui n'ont pas été détectés sur le champ, ont pu historiquement intervenir et se répercutent inévitablement dans les rattachements de données nationales qu'il est possible de requêter aujourd'hui.

Dans le cas favorable où l'unité d'expression du paramètre a été correctement renseignée, ce qui n'est pas toujours le cas, certaines valeurs de la même colonne sont de **vrais résultats quantifiés d'analyse**, d'autres des valeurs de **seuils analytiques de détection ou de quantification non atteints** (qui n'ont pas la même portée selon le paramètre et la méthode d'analyse en cause), qu'il faut ensuite impérativement post-traiter pour pouvoir ensuite calculer des variables statistiques (moyenne, médiane, quartiles...) concernant ce paramètre.

Dans l'état actuel, les dimensions et le volume de stockage accessibles dans des bases de données sont beaucoup moins contraignants qu'il y a 20 à 25 ans, mais l'inconvénient lié à cette structure de base, qui est restée figée depuis cette époque, est qu'à chaque nouvelle requête visant à utiliser les valeurs numériques de paramètres, la question du statut de la valeur inscrite dans la colonne de résultat, qui n'est pas forcément un résultat, se pose à nouveau avec les mêmes manques de données d'accompagnement (le plus souvent, lacune de code-remarque, parfois non-renseignement de l'unité) et il faut refaire de A à Z un post-traitement de données affectant une valeur numérique réelle et basse à toutes les cellules où le résultat analytique était inférieur à un seuil analytique. A noter que cette tâche est difficilement automatisable sans expertise et examen réel de données faisant intervenir un agent humain et une analyse attentive des diverses combinaisons de champs vides.

Ici, la première pratique a consisté, par des tris successifs sur les colonnes et les valeurs dans les cellules, à expertiser et vérifier la vraisemblance des données au même site ou pouvaient manquer unité et/ou code-remarque, et à compléter ce qui logiquement pouvait l'être. Ensuite, lorsqu'un seuil de détection ou un seuil de quantification n'était pas atteint (code-remarque adéquat et explicitement renseigné), il était affecté une valeur numérique faible, égale en général à 1/3 du SD ou du SQ attaché à la donnée pour la plupart des paramètres hydrochimiques naturels ou d'anthropisation, et ramenée à 1/2 du SD ou du SQ pour quelques paramètres ou cette pratique est apparue plus convenable pour le respect du gradient des basses valeurs.

2.4.3. Réalisation de l'agrégation des données disponibles

Rappelons que la demande européenne du GIG, en application de documents-guides Ecostat, portait sur une agrégation des données de chimie pour une expression en moyenne annuelle. Donc, même si la pertinence de cette approche est discutable pour le maillon des diatomées benthiques, dont la durée d'intégration biologique est bien inférieure à l'année, il convenait de la respecter pour participer à l'exercice.

Ce formatage particulier de données s'inscrit en cohérence avec une autre demande du GIG qui, afin d'étayer l'assise et la performance d'évaluation des systèmes nationaux soumis à l'intercalibration, sollicitait l'illustration, à la charge des Etats-Membres, de relations pression-impact entre maillon biologique et certains nutriments (P Tot, PO₄, NTD etc...) traités en moyenne annuelle.

Enfin, c'est sur la base des référentiels envoyés par chaque Etat-Membre qu'a été établie la relation pression-impact entre le gradient de PO₄ et l'évaluation diatomique réalisée selon l'étalon commun d'intercalibration diatomique (ou ICM diatomique) du jeu de données d'intercalibration ; cette relation pression-impact servant de passerelle permettant à l'exercice de Fit-in de se raccrocher à l'exercice précédent du X-GIG « Large Rivers » et d'établir la conformité des résultats obtenus dans le cadre du 2^{ème} exercice par rapport à ceux du premier, qui servent de référence.

Suite aux 2 séries d'opérations de complétion et de post-traitement de données qui précèdent, il a été mis rapidement en œuvre l'agrégation des résultats d'Opécont-Chimie disponibles en moyennes annuelles, selon la méthode d'application du pas de temps décrite en Figure 1, dans l'optique de la fourniture rapide du jeu de données d'intercalibration. A noter que la quasi-totalité des descripteurs de nutriments pour lesquels il existait de la chimie (hors années 2012 et 2013, ou les données de certains bassins n'étaient pas encore rapatriées) ont pu être renseignés de façon satisfaisante sous cette forme de moyennes annuelles. Dans ces conditions, les agrégations réalisées en été 2015 n'ont pas seulement concerné les 221 relevés nationaux transmis à l'U.E., mais ont concerné en tout, pour la réalisation de l'étude nationale, 523 relevés sur sites de plus de 8 000 km² (dont 433 concernant des sites de plus de 10 000 km²).

La requête complémentaire réalisée en Juillet 2017, trop tardive par rapport aux échéanciers opérationnels, n'a eu aucune répercussion sur l'exercice d'intercalibration UE. Par contre, elle a logiquement augmenté de façon significative (près de 20 %) l'assise de doublets de données mobilisables dans le cadre de l'étude nationale, qui a pu atteindre à ce stade 667 relevés sur sites de plus de 8 000 km² (dont 550 relatifs à des sites de plus de 10 000 km²).

Dans le cadre de ces 2 séries de requêtes successives, il a aussi été possible d'associer les données abiotiques calculées **sur une base trimestrielle** (selon le principe de requêtage indiqué en haut de la Figure 2) avec exactement le même nombre de relevés diatomiques que ci-dessus.

Les rapprochements évoqués ci-dessus ont pu être menés à bien, à partir de procédures assez simples de requêtage et de calcul, en ce qui concerne les **données de mesures physico-chimiques *in situ*** et les analyses relatives aux **teneurs des eaux en nutriments**.

Par contre, Il n'en a pas été de même en matière de référentiels représentatifs des **environnements hydrochimiques naturels** sous influence dominante du substrat géologique. En effet, l'assise-données collectée dans le cadre de nos réseaux naturels est nettement plus faible en cette matière (les analyses complètes étant réalisées de façon plus épisodique que les analyses régulières précédemment évoquées). Cependant, le GIG demandait une typification hydrochimique des TGCE nationaux soumis à l'intercalibration, sur la base d'un critère (alcalinité en moyenne annuelle) non-utilisé dans le cadre de nos réseaux de surveillance. Le chapitre suivant récapitule comment nous avons traité cette problématique spécifique et apporté une solution opérationnelle permettant de satisfaire à la demande de court-terme du GIG.

2.5. Post-traitement et complétion des données hydrochimiques spécifiques

2.5.1. Aspects généraux

En vue de réaliser les exercices d'intercalibration homogènes, portant sur des TGCE présentant des caractéristiques hydrochimiques cohérentes, le GIG « VLRs » a sollicité des données hydrochimiques naturelles attachées à chaque relevé et nous a demandé d'affecter chaque doublet [TGCE / site suivi] dans un des 2 types hydrochimiques européens suivants :

- Type R-L1 : cours d'eau de faible alcalinité (<0,5 meq CaCO₃ /l)
- Type R-L2 : cours d'eau alcalins (>0,5 meq CaCO₃ /l)

En France, malheureusement, l'alcalinité totale n'est pas un paramètre analytique fourni classiquement dans les résultats analytiques. De plus, même en utilisant les quelques analyses d'eau complètes effectuées, son calcul est impossible à reconstituer faute de dosage des ions hydroxydes.

Le paramètre analytique réalisé en France et le plus facile à rapprocher de l'alcalinité totale, moyennant la conversion nécessaire, est le TAC (Titre Alcalimétrique Complet) exprimé en degrés français. Le TAC est obtenu en mesurant le caractère alcalin de l'eau par titration avec de l'HCl jusqu'au virage acide des colorants hélianthine ou vert de bromocrésol (qui se produit dans les 2 cas dans une zone très proche du pH 3,2).

Pour pouvoir calculer un TAC moyen par site d'observation, condition nécessaire pour pouvoir approcher la valeur d'alcalinité demandée et affecter le site dans le bon type hydrochimique du GIG, il a fallu compléter de façon très importante les valeurs de TAC pour chaque relevé diatomique unitaire obtenu à un site, puis calculer ensuite le TAC moyen du site afin de pouvoir ensuite le situer dans la bonne typologie hydrochimique de l'exercice d'intercalibration.

Dans le jeu de données chimiques extrait en été 2015 pour préparer l'exercice d'intercalibration, pour mémoire, seulement **1 279 valeurs de TAC** ont été quantifiées sur les **7 772 Opécont-Chimie requêtées**, ce qui est très peu. En utilisant les données unitaires ou moyennées d'analyse de TAC disponibles en relation avec les relevés sélectionnés, seulement **255 relevés sur 645** ont pu être renseignés de façon directe.

Sur un plan plus large, les référentiels nationaux collectés sur les données hydrochimiques générales de l'eau, caractéristiques des environnements naturels, sont réalisées de façon nettement plus épisodique que les analyses physico-chimiques ou que les analyses de nutriments permettant de surveiller le niveau d'altération anthropique. En effet, il s'agit là d'analyses complètes de l'eau, plus coûteuses, donc réalisées nettement moins souvent que les autres séries d'analyses (en général, 4 analyses par an pour de 10 à 12 analyses par an sur les autres séries de paramètres).

Au-delà de la lacune relative en données de TAC, toute la série des paramètres hydrochimiques représentatives des environnements naturels se trouve moins souvent implémentée, donc forcément plus lacunaire que les chroniques d'autres types d'analyses.

Face à toutes ces données de TAC et d'autres paramètres manquantes, il a fallu par approches croisées, en se basant sur les valeurs d'autres paramètres analytiques qui lui sont relativement bien corrélés, produire une estimation approchée du TAC par modélisation (utilisation de régressions linéaires).

L'objectif principal de cette phase de travail était d'établir des relations solides entre les résultats analytiques de divers paramètres dans l'optique de la **reconstitution de données manquantes**, pas de représenter l'état moyen du site dans un créneau temporel.

Bien que plusieurs données de paramètres unitaires de l'eau (carbonates, calcium etc...) obtenues à des dates diverses aient pu être disponibles en correspondance avec un relevé diatomique donné, il faut ici préciser que, l'hydrochimie d'une eau s'établissant **dans un certain équilibre entre formes ioniques à une date donnée** et le but poursuivi étant la reconstitution la plus précise possible d'une donnée manquante à une date donnée, **il a été privilégié le ou les doublets de données obtenus à la même date pour établir ces régressions**. La donnée la plus globalisante (TAC) étant la moins souvent disponible et étant par ailleurs l'objectif principal de ces reconstitutions de données, vu que c'est ce descripteur qui permet d'approcher le plus facilement l'alcalinité totale demandée par le GIG, nous nous sommes calés sur les dates où le TAC était analysé pour établir les relations avec les autres descripteurs hydrochimiques de l'eau analysés à la même date.

2.5.2. Régression TAC / carbonates totaux

Bien que ne tenant pas compte des ions hydroxydes OH⁻, pour lesquels nous ne disposons pas de résultats analytiques directs, la concentration de l'eau en carbonates totaux [HCO₃⁻ + CO₃²⁻], qui contribue fortement à la charge alcaline des eaux naturelles, permet en général d'approcher très correctement les valeurs de TAC.

La relation présentée en **Figure 3** met en relation les valeurs de carbonates totaux et les valeurs de TAC obtenues sur des eaux prélevées au même lieu à la même date en association à un relevé biologique donné. Lorsque le TAC n'était analysé qu'une fois, les valeurs correspondantes des paramètres de la même analyse d'eau (ici, HCO₃⁻ et CO₃²⁻) ont été utilisées pour cette relation et sommées en concentrations équivalentes de CO₃²⁻. Lorsque 2 résultats ou plus de résultats de TAC étaient disponibles, il était procédé à la réalisation de moyennes sur le TAC et sur les carbonates totaux avec exactement la même assise de dates.

C'est selon cette façon de faire que les données existantes peuvent être reliées entre elles avec la meilleure qualité de relation. Et logiquement, la relation obtenue ainsi est certainement aussi la plus fiable pour la reconstitution d'une donnée manquante d'un doublet avec l'autre composante du doublet à la même date.

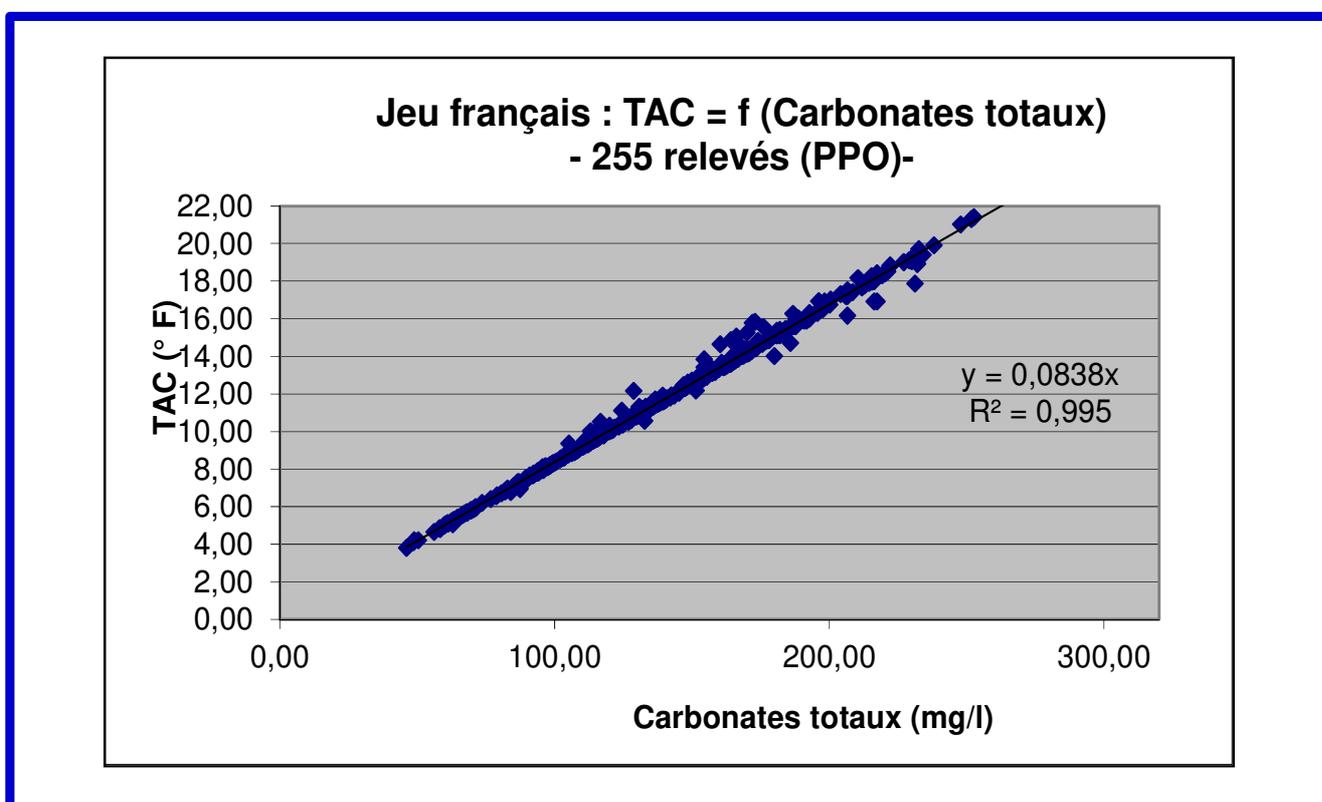


Figure 3 : Régression linéaire TAC /Carbonates totaux (255 doublets de données existantes)

Selon cette méthodologie d'association de données analytiques sur la même assise de date(s) de prélèvement, on obtient un excellent ajustement linéaire et un excellent coefficient de détermination entre ces 2 paramètres, sans aucun enlèvement d'outlier.

La **Figure 4** en page suivante illustre, par extension, la même relation après incorporation des données manquantes reconstituées sur la base de l'ajustement linéaire ci-dessous reporté (i.e. équation d'ajustement linéaire obtenue en Figure 3) :

$$y = 0,0838 x$$

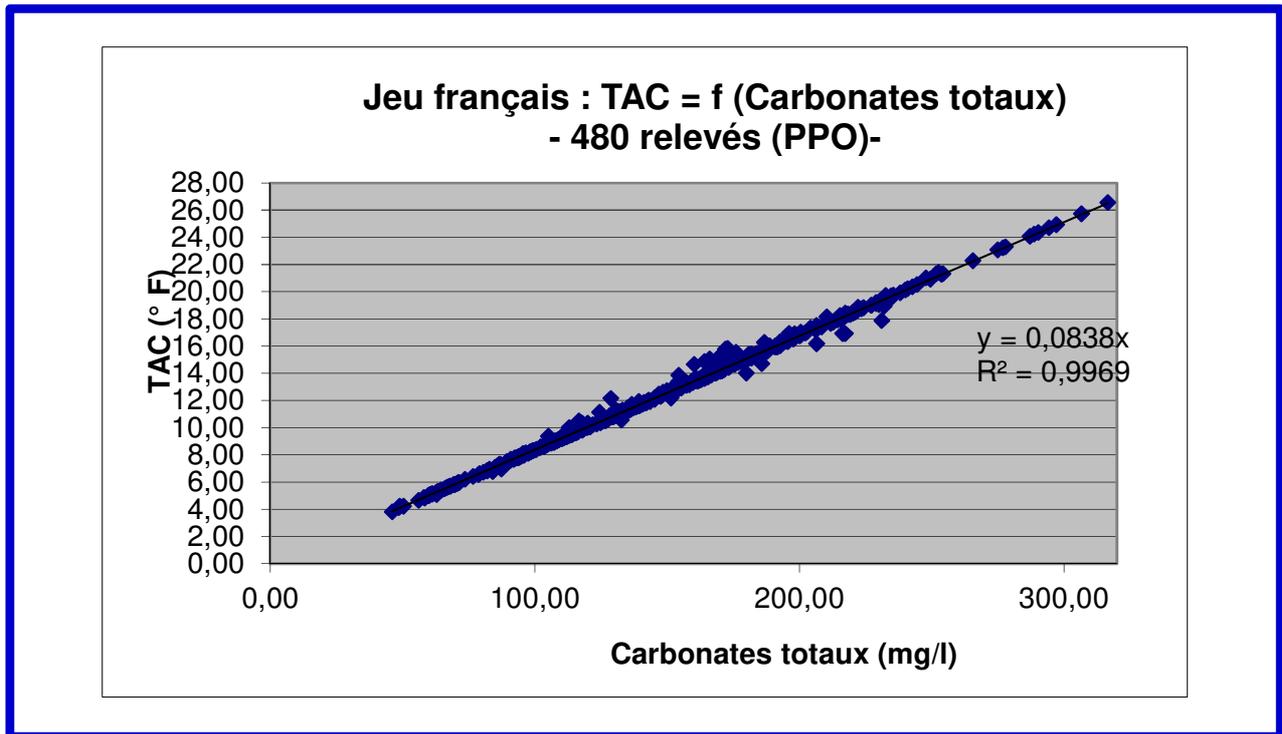


Figure 4 : Reconstitution de données manquantes à l'aide de la régression TAC /Carbonates totaux

Sur la base des 255 doublets de données initialement disponibles et grâce à la très bonne qualité de la régression qui en a résulté, les données de TAC attachées à des relevés diatomiques individuels ont ainsi pu être complétées dans un nombre très substantiel de cas, permettant presque un doublement du nombre de relevés renseignés sur ce paramètre. Cette complétion a pu intervenir dans les 2 sens :

- dans de rares cas, expertise et/ou complétion des données de carbonates à partir d'un titrage de TAC fait à la même date (complétion de 18 relevés).
- reconstitution de données de TAC manquantes à partir de données de carbonates existantes, qui était de loin le cas le plus fréquent (complétion de 207 relevés).

Au bilan, cette étude de régression entre Carbonates totaux et TAC basée sur **255** doublets de données analytiques réelles a déjà permis presque un doublement des données de chimie complètes (**225** autres relevés ont pu être complétés de cette façon), portant à **480** le nombre de relevés avec données complètes et rendant possible une classification satisfaisante de ces relevés diatomiques selon les types hydrochimiques définis par le GIG.

2.5.3. Régression TAC / Ca⁺⁺

Le TAC est un paramètre souvent assez bien relié à la quantité de cations Ca⁺⁺ ou (Ca⁺⁺ + Mg⁺⁺), rendant *a priori* intéressant l'examen de cette relation.

Son résultat en est donc illustré à titre indicatif dans la **Figure 5** page suivante. Cette relation ne pouvait permettre de compléter qu'une petite dizaine de données manquantes de TAC. Or on voit bien que le jeu de données national englobe plusieurs typologies de relations hydrochimiques dont 2 très différentes et une intermédiaire, selon les caractéristiques géologiques et hydrochimiques des bassins versants.

Un nuage principal concerne une très nette majorité des sites Français surtout concernés par le gradient carbonaté « normal » issu de placages sédimentaires calcaires (ex : Bassin Parisien, Coteaux Aquitains etc...).

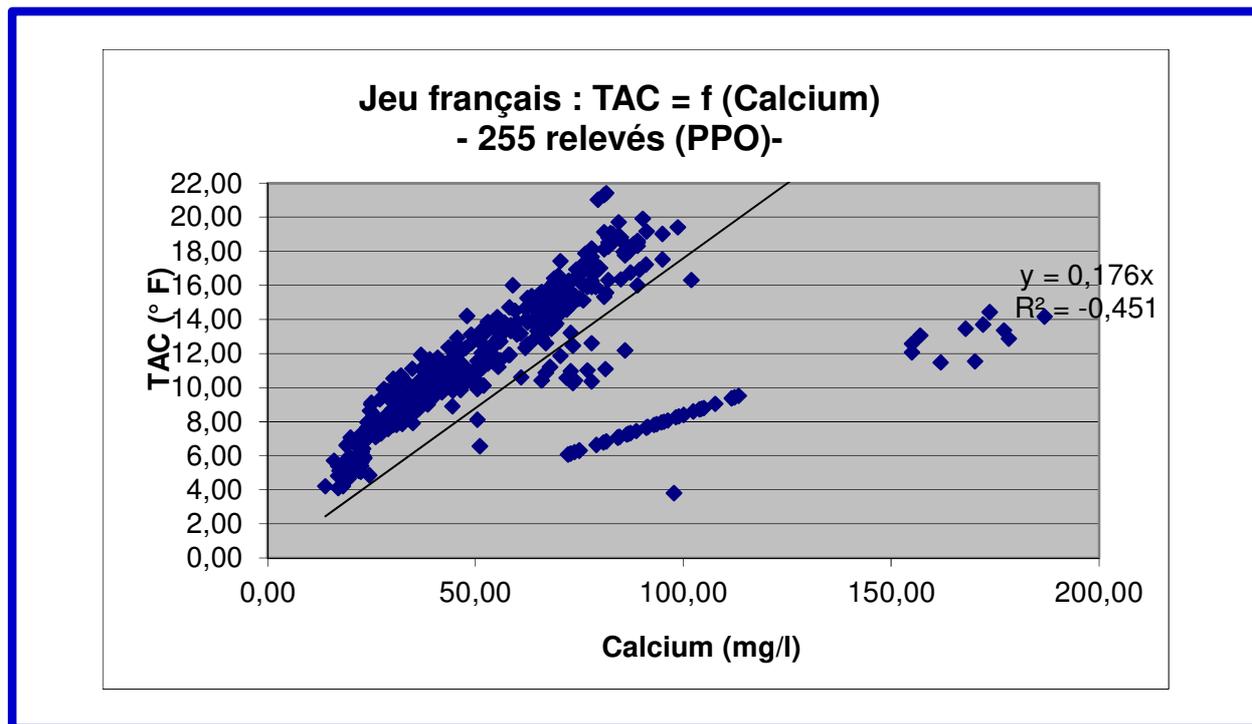


Figure 5 : Relation entre calcium et TAC (255 doublets de données existantes)

En complément de ce nuage principal, il a aussi été possible de repérer et de qualifier des sites et cours d'eau à particularités hydrochimiques :

- 2 points exceptionnels tout en haut à droite (Saône à St Symphorien d'Annelles, à alcalinité exceptionnellement forte par rapport à la teneur en Ca⁺⁺). La Saône est réputée subir des influences salines parfois naturelles, mais aussi industrielles (notamment usines Solvay de Tavaux). Ce point situé quelques kms à l'amont de l'entrée dans la communauté de communes du grand Lyon se situe sous l'influence amont de plusieurs villes relativement importantes (Chalon, Mâcon, Tournus, Villefranche-sur-Saône) et subit sans doute l'influence de déversements de substances chimiques basiques (NaOH, KOH ...)
- 1 nuage à positionnement particulier en bas à droite (Moselle). Là, le TAC est plutôt modéré par rapport à la teneur en calcium, mais nous sommes en présence d'un contexte minéralogique aux particularités bien connues, qui entraîne une hydrochimie très minéralisée et chargée en chlorure, en sodium et en sulfates ;
- 1 alignement Seine-Marne-Oise à faible TAC par rapport à la charge en calcium, qui s'aligne avec le cas précédent de la Moselle ;
- Et entre les 2 alignements principaux précités, des rivières arrivant de zones montagneuses peu à modérément carbonatées (valeurs de TAC plus basses que pour le nuage principal) qui correspondent à la Dordogne et à l'Isère.

En raison du faible enjeu de complétion sous-jacent, il n'a pas été tenté d'isoler les groupes de relevés afin d'obtenir 3 régressions plus pures correspondant à des groupes de cours d'eau plus comparables dans leur équilibre hydrochimique. C'était bien sûr faisable et aurait certainement permis un bon gain d'ajustement des régressions résultantes, mais aurait été trop faiblement valorisé vu le faible nombre de points potentiellement concernés.

Cette étude de régression a cependant donné une occasion intéressante de bien repérer et passer en revue les spécificités et les regroupements hydrochimiques possibles de nos TGCE Français sur le plan naturel, en fonction de l'influence hydrochimique qui découle des caractéristiques géologiques et pédologiques de leurs bassins versants. Cette visualisation des données aide à mettre en relief certaines différences typologiques naturelles dans notre ensemble national de très grands cours d'eau, ce qui constitue une trame de fond intéressante dans le cadre de l'étude nationale visant à étudier et à proposer un nouveau système d'évaluation de nos TGCE.

2.5.4. Régression TAC / [Cond. El. * pH]

L'idée de cette agrégation de paramètres sous forme d'un produit est venue du fait que le résultat de leur combinaison est pertinent pour rendre compte d'une gradation sur le plan de la minéralisation et de l'alcalinité de l'eau.

La conductivité électrique est fortement fonction de la charge ionique de l'eau. Toutefois, même si, dans les conditions naturelles des milieux aquatiques d'eau douce, les plus fortes conductivités sont le plus souvent liées à des conditions hydrochimiques plutôt alcalines (substrats géologiques calco-magnésiens) sous influence d'ions hydrogénocarbonates, carbonates et hydroxydes, ce peut ne pas être toujours le cas. Certaines sources thermales avec forte charge ionique peuvent être acides et réductrices (cas des sources d'origine tellurique à basses valeurs de redox, en liaison avec l'émergence de cations acides (H₃O⁺), de sulfures...). Dans un contexte plus anthropisé, des eaux chargées d'acides émanant par exemple d'industries chimiques peuvent aussi être doté d'une conductivité électrique moyenne à forte tout en étant très acides (s'accompagnant de valeurs basses de pH dans ce cas). Cette dernière situation est néanmoins en voie de raréfaction et ne peut pas rester chronique du fait de la réglementation actuelle sur les pollutions ponctuelles et industrielles.

Pour le pH, la liaison des hautes valeurs avec la présence d'espèces chimiques alcalinisantes (ions hydrogénocarbonates et carbonates, ions hydroxydes) est fréquente mais pas absolue. En effet, dans des eaux peu minéralisées, le pH est très facilement variable du fait du faible pouvoir-tampon des eaux, pouvant facilement passer d'une valeur acide en cas de présence d'une quantité modeste d'ions acides à une valeur montant rapidement et fortement jusqu'à des valeurs élevées, en liaison par exemple avec la photosynthèse et une sursaturation de l'eau en O₂.

Par contre, en conditions fortement minéralisées (conductivités électriques moyennes à fortes), l'expression d'un pH au-dessus de la neutralité voire basique va plutôt traduire les caractéristiques hydrochimiques alcalinisantes qui siègent dans la colonne d'eau.

Le produit de ces 2 paramètres a donc du sens pour établir un proxy de l'alcalinité de l'eau car l'indication de la valeur de pH, assez versatile dans les conditions peu minéralisées et avec une forte influence confondante de la photosynthèse, est plus stabilisée en cas de forte conductivité, du fait du pouvoir-tampon de l'eau, que ce pH soit dans le domaine acide ou au contraire dans le domaine basique.

La mise en relation brute entre (pH x Cond. El.) et TAC à partir des valeurs disponibles en doublet sur 483 relevés a permis l'expression de la relation présentée en **Figure 6** ci-dessous.

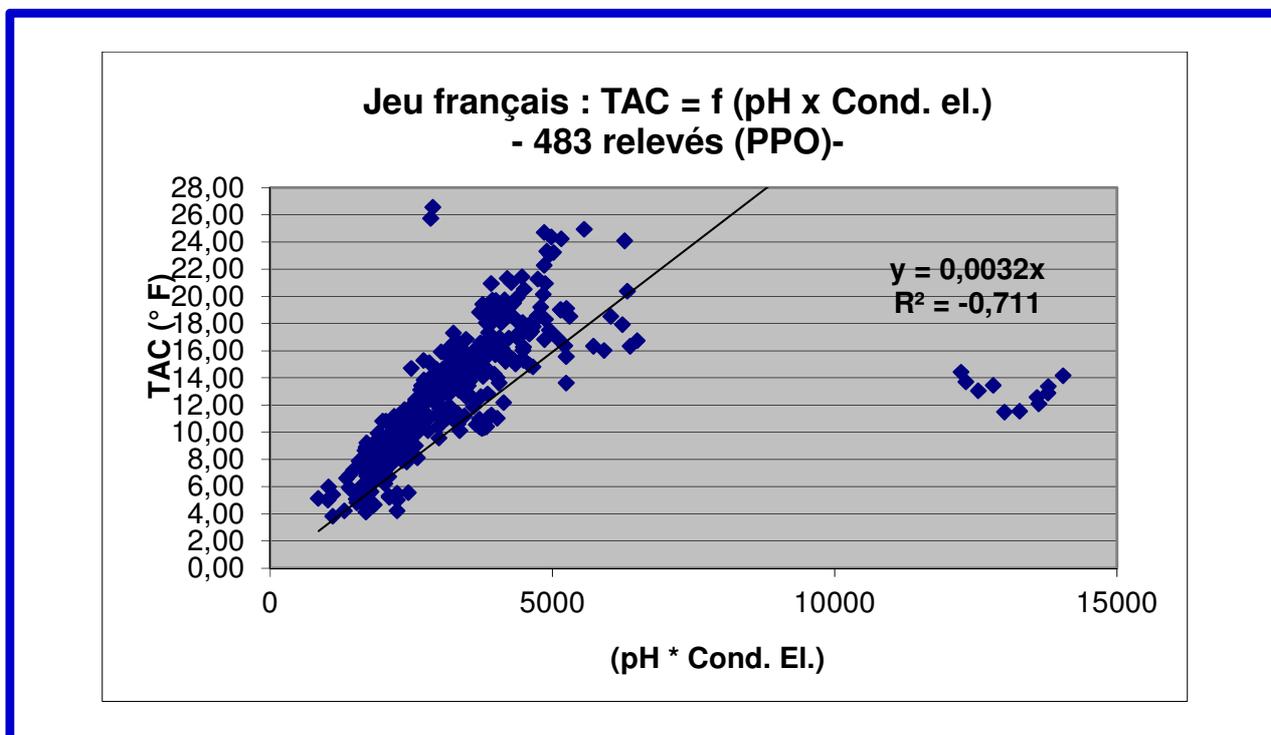


Figure 6 : Régression linéaire tout-venant entre (pH * Cond. El.) et TAC (sur 483 doublets de données existantes)

Sur ce premier essai ou le jeu de données national TGCE est traité globalement, l'examen du résultat obtenu avec la combinaison de ces 2 variables complémentaires sous forme d'un produit [pH * Cond. El.], l'une indiquant un positionnement entre domaine acide – domaine basique et l'autre rendant surtout compte d'un degré de minéralisation de l'eau, rend visiblement mieux compte du domaine hydrochimique auquel appartient une eau que ces 2 variables isolées.

En effet, il est patent qu'une relation à validité quasi-générale ressort, mais que 2 groupes d'exceptions typologiques, très éloignés de l'ajustement dominant pour les TGCE Français, mettent à mal le R^2 (valeur négative, dénuée de significativité), alors que la relation entre les 2 descripteurs est positive et semble être chargée de sens.

Les paramètres pH et conductivité électrique étant disponibles, souvent à plusieurs dates, en couplage avec tous les relevés diatomiques, il a donc été jugé intéressant de pousser un peu plus loin et de tenter de s'appuyer sur leur combinaison en vue de compléter sélectivement le maximum de valeurs de TAC encore manquantes.

Comme un nombre encore assez conséquent de valeurs de TAC (au moins une cinquantaine) sont encore manquantes et gagneraient à être complétées autant que possible, sous réserve de l'existence de données de chimie correspondantes, il a été décidé de tenter de mieux spécifier cette relation vis-à-vis des types de cours d'eau représentés, en sortant les exceptions typologiques et en repérant bien les cours d'eau sur lesquels s'applique un ajustement linéaire différent (auquel cas, il ne serait pas valide d'inférer la valeur de TAC manquante).

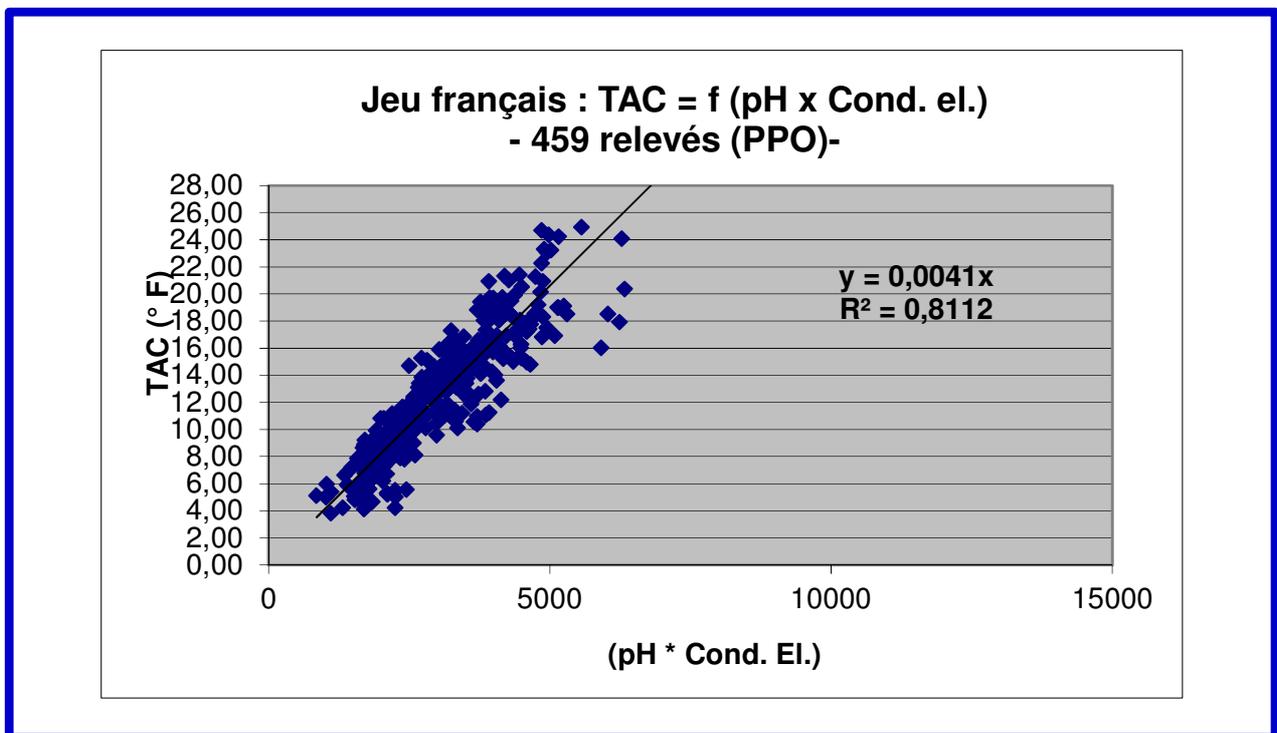


Figure 7 : Régression linéaire entre (pH * Cond. El.) et TAC
(sur une sélection de 459 doublets de données existantes)

La **Figure 7** illustre qu'en enlevant les 24 relevés les plus éloignés de l'ajustement principal, soit environ 5 % correspondant à des typologies hydrochimiques sensiblement différentes, la relation principale s'en trouve déjà très fortement confortée, avec une augmentation sensible de la pente de la relation et un R^2 que l'on peut désormais qualifier de satisfaisant (0,81), notamment en vue de procéder à une complétion de données.

Les 26 relevés écartés de la typologie dominante concernent des rivières franchement décalées par leurs relations hydrochimiques qui, en cas de donnée de TAC manquante, ne peuvent pas être complétées sur la base de cet ajustement linéaire. Il s'agit de :

- 1) 14 relevés correspondant à une typologie très spécifique (rivières salées), qui concerne la Moselle (11 relevés) et la Saône à Seurre (3 relevés), au début de son parcours de TGCE. Il s'agit du paquet de relevés qui, dans la Figure 5, se situait complètement en bas à droite du nuage général ;

- 2) le Rhin à Lauterbourg (2 relevés aux valeurs les plus élevées de TAC, qui avoisinent ou dépassent 26 Degrés Français) ;
- 3) la Durance à Vinon (3), l'Isère à Tullins et St Gervais (5 relevés / 12) charge minérale importante mais TAC relativement faible.

Pour consolider encore la relation du nuage principal de points, ce qui peut encore améliorer les choses pour la reconstitution de données manquantes appartenant au nuage principal, il a été tenté de supprimer encore quelques points les plus en marge de cette relation. La plupart du temps, l'explication ne tient plus à une rivière se rattachant à une typologie spécifique, entièrement ou en un lieu, mais à une date exceptionnelle, ou à un petit nombre de dates exceptionnelles dans une série temporelle de relevés pour lesquelles, pour des raisons ayant essentiellement trait à l'hydrologie et à son influence sur des lois de dilution et d'équilibre, des relevés de certaines rivières se trouvent dans des valeurs de TAC plus faibles que ne le laisserait présager le rapport [pH * Cond. El.].

Il a donc été procédé pas à pas à l'enlèvement de valeurs les plus décalées de la relation avec repérage des rivières concernées. 15 relevés supplémentaires les plus décalés de la relation, représentant un pourcentage de 3,25% de valeurs dites "outliers", ont aussi été mises de côté du jeu de données à ce stade. La relation la plus consolidée pour la reconstitution de données de TAC sur des cours d'eau appartenant à la typologie dominante s'établit désormais sur 444 doublets de données (voir **Figure 8** ci-dessous).

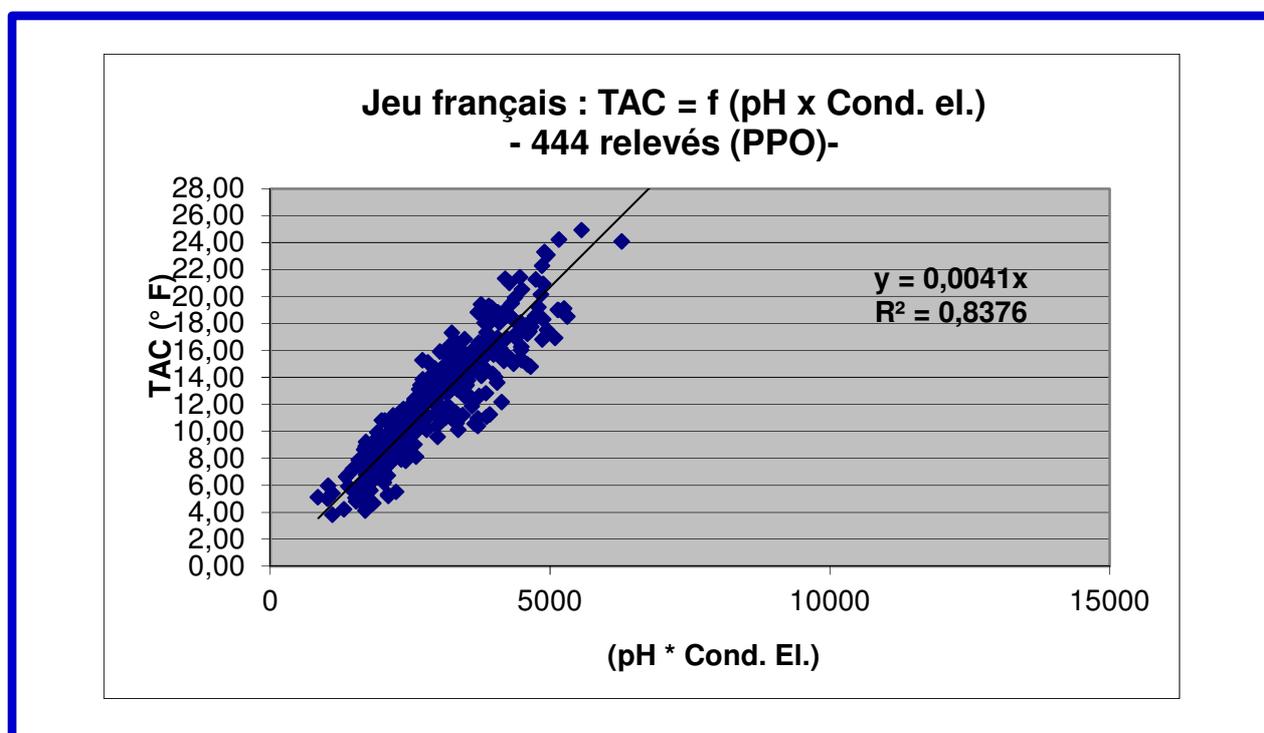


Figure 8 : Régression linéaire consolidée entre [pH * Cond. El.] et TAC
(sur une sélection de 444 doublets de données existantes)

Les 15 relevés "outliers" les plus décalés de la relation concernaient la plupart du temps des cours d'eau à particularités hydrochimiques bien connues, soit peu minéralisées et en-dessous du nuage dominant (la Vilaine à St Dolay, l'Oise à Compiègne), soit fortement minéralisée et au sommet du nuage, comme la Garonne à Port Ste Marie (2 relevés).

Au total, que la raison ait pour origine un cours d'eau en exception typologique d'origine géochimique, ou une hydrologie un peu exceptionnelle conduisant à une valeur moyenne annuelle un peu atypique pour le cours d'eau en ce lieu, les valeurs de 39 relevés sortant manifestement de la typologie dominante, soit environ 8% de relevés, ont été sortis de l'assise de cette régression.

La loi de régression qui en résulte convient pour la reconstitution de valeurs de TAC sur la plupart des cours d'eau de France, hormis pour les quelques rivières repérées comme différentes sur le plan typologique, ou certains sites très spécifiques sur les rivières précitées (Moselle, Saône à Seurre, Rhin à Lauterbourg, Durance à Vinon et Isère dans le secteur de Tullins-St Gervais).

Pour des manques de données de TAC concernant ces cours d'eau et sites spécifiques, la régression donnée en Figure 8 n'est pas convenable et ne peut pas s'appliquer valablement. Il vaut donc mieux se référer aux autres données obtenues pour reconstituer une donnée manquante sur ce même cours d'eau à ce site, et c'est ainsi qu'il a été procédé dans les rares relevés où le cas se posait .

Après ré-inclusion des 41 nouvelles données de TAC modélisées, le nuage complet de points obtenu présente la physionomie illustrée en **Figure 9** ci-dessous. Elle peut se comparer directement au nuage initial de points présenté à la Figure 6 page 19. Sauf pour 2 points à compléter, repérés en exceptions typologiques, les points rajoutés l'ont été selon la relation linéaire consolidée présentée en Figure 8 (R^2 proche de 0,84).

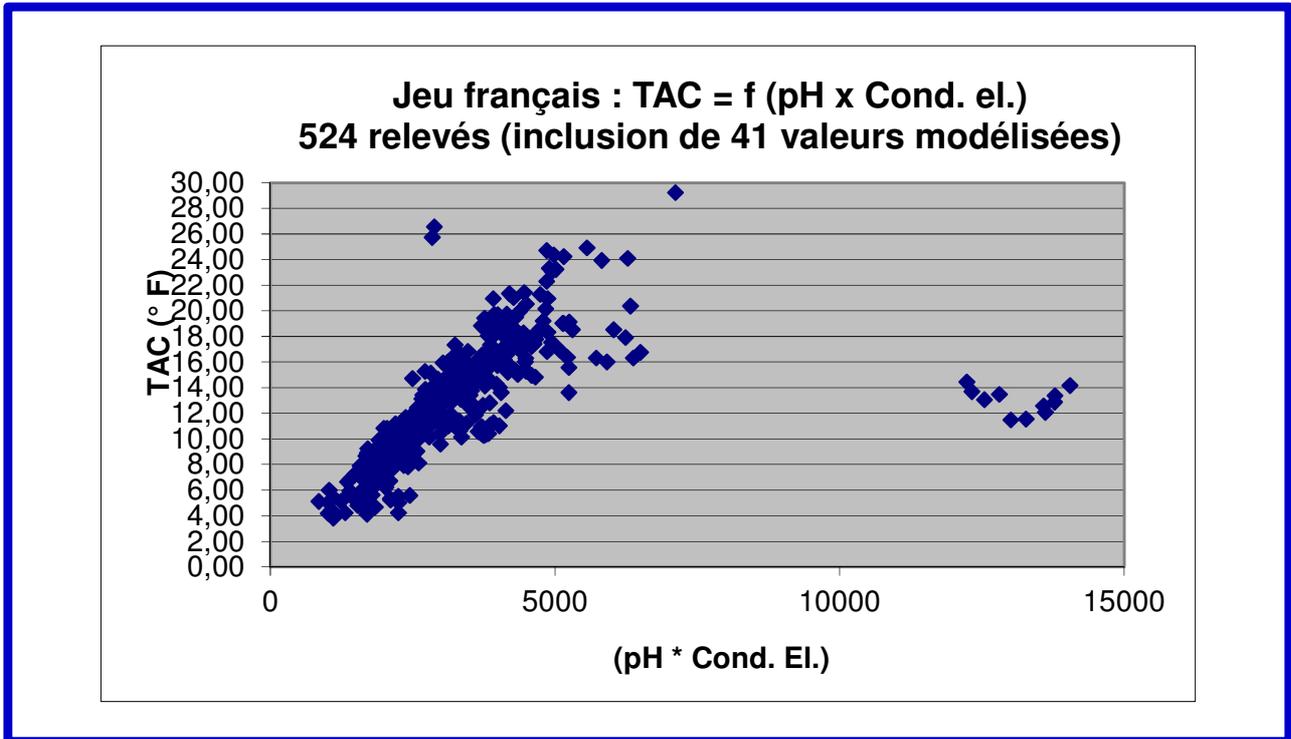


Figure 9 : Jeu de données complété sur la base de la relation $[TAC = 0,0041 (pH * Cond.El)]$ (524 doublets de données)

Au bilan, donc, ces opérations ciblées de reconstitution de données de TAC manquantes ont permis de plus que doubler les relevés renseignés vis-à-vis de ce paramètre (passage de 255 à 524 relevés renseignés), permettant une nette consolidation de la classification de sites selon les 2 types de VLRs définis par la typologie européenne :

- 1) le **Type R-L1** : cours d'eau de faible alcalinité ($<0,5 \text{ meq CO}_3^{2-} / \text{l}$)
- 2) le **Type R-L2** : cours d'eau alcalins ($>0,5 \text{ meq CO}_3^{2-} / \text{l}$).

A noter que l'effectif total de 524 relevés renseignés pour la réalisation de l'étude nationale couvre tous les cours d'eau-candidats TGCE de plus de 8 000 km² de bassin versant. Ceux entre 8 000 et 10 000 km² ont donc participé aussi à l'assise de l'étude et ont fait l'objet d'une prise en compte optionnelle dans le cadre de la proposition du nouveau système national d'évaluation des TGCE.

Les données envoyées au GIG ont fait l'objet d'une sélection visant à donner une vision la plus représentative possible des TGCE français, à partir de l'envoi d'un jeu de données de 221 relevés strictement prélevés sur des BV intégrés de plus de 10 000 km², en conformité avec la typologie européenne en vigueur.

2.6. Calculs d'alcalinité au site

En vue de classer correctement l'hydrochimie des couples [relevés X sites] de notre jeu de données national soumis à intercalibration, dans le respect des critères requis par le GIG pour la participation à l'exercice, qui sollicitait une classification basée sur l'alcalinité exprimée en meq/l d'équivalent- CO_3 , nous avons procédé selon la séquence suivante :

- complétion des données de TAC, basée sur les lois de régressions linéaires évoquées aux chapitres précédents,
- re-calcul d'alcalinité (basé sur une moyenne annuelle au site) dans le système d'unité requis.

La conversion entre TAC en degrés Français et meq/l de CO_3^{2-} a été obtenue de la façon suivante :

<i>Alcalinité totale (exprimée en meq/l de CO_3^{2-}) = TAC /10 (en Degrés Français)</i>
--

Alors que ce n'est pas forcément le cas lorsque l'on considère des très petits et petits cours d'eau, qui peuvent plus facilement être issus d'un bassin versant à caractéristiques géologiques et hydrochimiques homogènes, le fait de considérer des très grands cours d'eau amène inévitablement à une origine souvent composite des eaux intégrées à cette échelle, certains petits affluents venant de zones carbonatées augmentant rapidement l'alcalinité et le pouvoir-tampon d'eaux peu minéralisées provenant de contextes géologiques plus acides (par exemple, massif hercynien).

Dans notre contexte Français, vu la mixité des géochimies de bassins versants à l'échelle de taille caractéristique des TGCE, il n'est pratiquement plus possible de trouver des TGCE à bassin versant complet faiblement carbonaté, pouvant être qualifiés de peu minéralisés. Au bilan, une seule rivière de plus de 10 000 km² à un seul site (Vilaine à St Dolay) a pu rentrer au final dans la catégorie des très grands cours d'eau peu minéralisés selon les critères typologiques du GIG.

Si l'on considère le seuil de taille optionnel de 8 000 km², bien que la mixité géochimique des bassins versants soit légèrement diminuée, seuls 4 hydrosystèmes, selon les stations, se rattachent aussi à ce type peu minéralisé (outre la Vilaine sur tout son parcours au-dessus de 8 000 km², la Dordogne entre 8000 et 10 000 km², ainsi que l'Allier et la Loire à l'amont de leur confluence).

2.7. Bilan des données post-traitées, sélection du jeu national envoyé au GIG

Le bilan des opérations effectuées successivement en été-automne 2015, puis en Juillet 2017 sur l'extraction actualisée de toutes les données TGCE disponibles (cf. Tableau 1 pour le bilan quantitatif global des données), peut donc être résumé de la façon suivante :

- 1) **Extraction des relevés diatomiques disponibles dans la base de données nationales Pandore**, sur une base temporelle garantissant une certaine homogénéité de l'indice, du guide taxonomique et des relevés biologiques (période retenue : entre 2007 et 2013 inclus).

Résultats : **645 relevés** diatomiques TGCE extraits selon un seuillage de surface minimale de BV de 8 000 km² (dont **532 relevés** sur un BV de plus de 10 000 km²). .)

- 2) **Extraction et complétion des données informatives manquantes sur les sites concernés par ces relevés** (coordonnées géographiques, HER, type national, numéro de masse d'eau, QMNA5, surface intégrée à la station de réseau etc...)

Résultats : **154 sites TGCE** initialement extraits selon un seuillage de surface minimale de BV de plus de 8 000 km², dont **99 sites de plus de 10 000 km²** répondant aux critères typologiques du GIG. En fonction de la complétion des données de chimie, la sélection des relevés biologiques incorporés dans l'envoi du jeu de données national en vue de la réalisation effective de l'exercice d'intercalibration a concerné en tout **95 sites nationaux de plus de 10 000 km²**.

- 3) **Extraction des Opécont-Chimie correspondant à :**

(date du relevé diatomique – 1 an) ⇒ (date du relevé diatomique + 15 jours)

Résultats : Extraction de **plus de 11 000 Opécont-Chimie différentes** couvrant de 2007 à fin 2013. Selon le cas, les Opécont-Chimie étaient complètes (assez rarement, incluant paramètres hydrochimiques naturels et nutriments), partielles (plus fréquentes, seulement nutriments), ou seulement physico-chimiques (souvent mensuelles, mesures in situ).

4) **Fusion / complétion des Opécont-Chimie** réalisées à la même date, résolution des problèmes des basses valeurs inférieures au seuil de quantification (SQ) ou au seuil de détection (SD), des problèmes d'unités, des problèmes de codes-remarques etc...).

5) **Calcul des moyennes annuelles**, à la demande du GIG, pour tous les paramètres ou c'était possible

Résultats : Sur **645 relevés diatomiques disponibles** (BV > 8 000 km²), **523 relevés sur sites de plus de 8 000 km²** ont des données de chimie complètes ou qui ont pu être complétées, dont **433 relevés sur sites de plus de 10 000 km²**.

6) Calcul de moyennes trimestrielles, dans le cadre de la réalisation de l'étude nationale « Elaboration d'un système d'évaluation des TGCE de France », selon le principe présenté ci-dessous :

(date du relevé diatomique – 75 jours) ⇒ (date du relevé diatomique + 15 jours)

Résultats : Idem alinéa 5 ci-dessus.

7) **Complétion des données naturelles de géochimie-hydrochimie manquantes** par relevé biologique, en vue de conforter la classification typologique de nos TGCE dans un système conforme à la typologie européenne d'intercalibration (utilisation de l'alcalinité de l'eau, exprimée en meq CO₃/L). Complétion des données manquantes sur la base de modèles d'ajustement linéaire, puis intégration moyenne annuelle au site.

Résultats : passage de 255 à 524 relevés biologiques complétés sur le plan de l'hydrochimie naturelle.

8) **Sélection et envoi des jeux de données d'intercalibration TGCE au GIG VLR.** Les données envoyées ont fait l'objet d'une sélection finale basée : 1) sur la complétion effective de toutes les données demandées dans le fichier-type (template) européen ; 2) en essayant de représenter le gradient maximum de situations écologiques rencontré à chaque site sur la chronique temporelle sélectionnée (2007-2013). Lorsque possible, 3 relevés ont été sélectionnés au même site : celui ayant obtenu la meilleure note d'indice, celui ayant obtenu la plus basse note d'indice et un relevé ayant obtenu une notation intermédiaire. Dans la mesure où ce qui est inter-comparé et intercalibré avec l'étalon européen (ICM) sont des évaluations obtenues au relevé biologique grâce à une notation indicielle normée en EQRs par région naturelle, cette façon de sélectionner les données a permis une meilleure représentation des gradients naturels et d'altération possibles à un même site selon l'année et le niveau spécifique des pressions anthropiques reliées au relevé biologique. En effet, l'important dans l'exercice collectif réalisé n'était pas de se centrer sur la classification écologique du site et sa variabilité, ou sur son état moyen, mais sur l'aptitude du système national d'évaluation pour classer l'état écologique au relevé d'une façon comparable au système-étalon européen, si possible en toutes situations (hydrologiques, de pollution...). Cette façon de procéder pour la sélection finale des données est celle qui apporte la meilleure garantie de bon étalonnage de notre système national, car elle permet de s'appuyer sur un gradient de situations aussi étendu que possible.

Résultats : Pour une réalisation équilibrée de l'exercice en inter-Pays-Membres, il était inutile et même gênant d'envoyer un nombre trop important de données. La sélection faite selon les principes décrits précédemment a permis, à partir des 433 relevés biologiques complets disponibles à cette époque sur TGCE > 10 000 km², d'envoyer un jeu de données national de **221 relevés**.

9) **Utilisation des 667 relevés issus de sites de plus de 8 000 km² (dont 550 relevés biologiques complets issus de sites de plus de 10 000 km²) pour la réalisation de l'étude nationale.** Cette assise complète de données a été utilisée à la fois pour le calcul des relations pressions-impacts à l'échelle de tous les TGCE de France, et pour venir en appui de la comparaison des statistiques d'évaluation procurées par 4 scénarios différents de déclinaison nationale possible du dispositif d'évaluation.

3. ETUDE DES RELATIONS PRESSION-IMPACT BASEES SUR L'IBD₂₀₀₇

3.1. Aspects généraux, éléments de contexte

Suivant en cela des requêtes formulées dans des documents-guides d'accompagnement à la réalisation des exercices d'intercalibration (nommés « Guidances ECOSTAT »), la coordinatrice du 2ème exercice européen d'intercalibration « Large Rivers » auquel la France a participé (Dr Dasa HLUBIKOVA, de nationalité Bulgare et travaillant à l'époque dans un Laboratoire de Recherche Autrichien), a sollicité de la part des Etats-Membres participants tout élément susceptible de venir étayer la description et la performance de leurs systèmes d'évaluation nationaux soumis à l'intercalibration. En contexte d'échéancier de réponse tendu, nous avons tout d'abord fourni des **éléments généraux de description du système d'évaluation français** tel que mis en œuvre au niveau national (cf. REF. 4). Il a aussi été fourni des éléments de performance en matière de réponse pression-impact (cf. REF. 5) produits en 2006-2007 pour venir étayer la mise en place du nouveau système d'évaluation basé sur l'IBD 2007 à l'échelle complète du territoire Français, **toutes tailles et tous types de cours d'eau confondus**.

Sur la base d'une 1^{ère} requête de Juin 2015, nous avons ensuite étudié des **relations pression-impact plus spécifiquement centrées sur des jeux de données TGCE**. Ces relations ont été établies dans un premier temps sur la base des données de chimie moyennées à l'année, comme demandé par le GIG. Les relations étant moins convaincantes que sur les analyses tous cours d'eau confondus, une ré-actualisation de ce travail a été tentée ultérieurement à partir d'une dernière actualisation du jeu de données extrait de PANDORE (requête de Juillet 2017), d'une part sur les moyennes annuelles de chimie, mais aussi sur un pas de temps plus court (intégration de 3 mois de chimie des eaux), afin d'étudier les réponses biologiques selon un créneau temporel plus représentatif de l'exposition subie par les biofilms diatomiques. Bien qu'une amélioration ait été notée, la démonstration de ce type de relations est moins évidente sur les grands hydrosystèmes que sur des cours d'eau de plus petite taille, pour différentes séries de raisons discutées dans les paragraphes suivants.

3.2. Comparaison des réponses de l'IBD2007 et de l'IPS :

A l'échelle générale de la France, au moment de la genèse de ce nouvel indice, nous avons évalué la performance biologique de la nouvelle version d'IBD (IBD₂₀₀₇), tous types de cours d'eau confondus, par rapport à la version antérieure utilisée à l'époque (IBD₂₀₀₀), mais aussi par rapport à l'IPS (REF. 6, Coste, M., 1982), qui est considéré comme un indice diatomique de référence au niveau européen et qui fait partie intégrante de l'ICM diatomique (Intercalibration Common Metric) adopté au niveau de l'Europe (cf. Figure 10 ci-dessous).

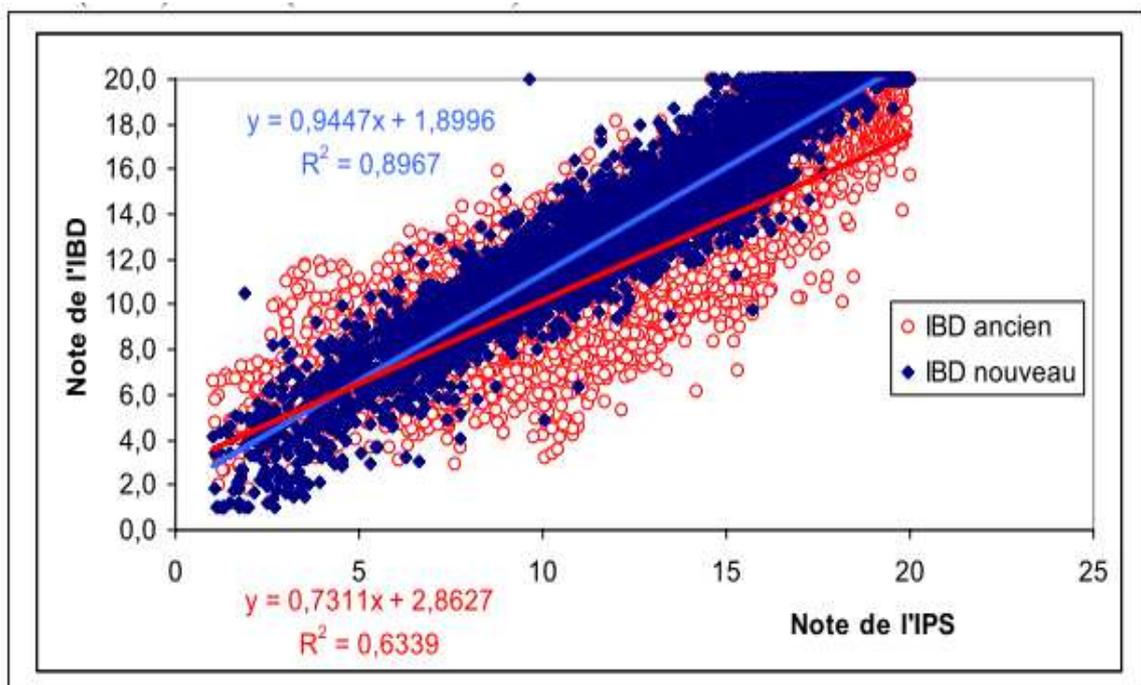


Figure 10: Régressions linéaires entre les 2 versions successives de l'IBD et l'IPS, indice diatomique de référence

Cette comparaison a été réalisée sur un jeu de données national historique constitué de 7 722 inventaires diatomiques disponibles en 2007. On peut y observer que la relation de l'IBD₂₀₀₇ (en bleu) avec l'IPS s'est améliorée par rapport à la version antérieure IBD₂₀₀₀ (en rouge), la progression se matérialisant à la fois par l'augmentation de la pente et par l'amélioration du coefficient de détermination (R^2), qui est désormais très proche de 0,9. Il est à préciser que cette relation a été établie sans aucun enlèvement préalable d'outlier(s).

3.3. Relations pression-impact entre IBD2007 et nutriments (tous types de cours d'eau confondus)

Les systèmes d'évaluation présentés par les Pays-Membres pouvant comporter d'importantes différences de principes d'évaluation et de philosophies sous-jacentes, la coordinatrice a tout d'abord demandé des informations sur le mode de construction et sur le principe de calcul de l'outil national présenté à l'intercalibration. Dans le cadre d'une note technique envoyée par la France (cf. REF. 4), il a donc été rappelé que l'IBD₂₀₀₇ (REF. 7, Coste et al, 2009) a été construit à partir d'un gradient composite d'anthropisation basé sur 7 descripteurs résumant différents types de pression anthropique (dans un ordre décroissant de contribution, NH₄ et PO₄, DBO₅, Sat. O₂, Cond. El., pH, NO₃).

L'indication du degré d'altération restitué par cet indice a été finement calée sur sa **réponse composite** à ces descripteurs. Il est bien sûr possible d'évaluer sa réponse pression-impact face à un seul descripteur de pression anthropique (par exemple PO₄, DBO₅...) pris en compte successivement. Mais il peut aussi être considéré comme logique que sa réponse biologique, construite sous l'influence intégrée des 7 paramètres précités, ne soit pas en parfaite concordance ou ne s'exprime pas forcément à son optimum vis-à-vis de l'effet séparé de chaque pression. Quoiqu'il en soit, pour répondre à la demande formulée, nous avons présenté 2 exemples de relations pression-impact mono-nutriment qui avaient été produites en 2007 en appui du changement de version de notre indice diatomique national (remplacement de l'IBD₂₀₀₀ par l'IBD₂₀₀₇).

Précisons que tous les échantillonnages biologiques pris en compte ici ont été collectés pendant la période de bas débits de la belle saison (soit entre Juin et Septembre) dans tous les types de rivières Françaises allant des plus petits cours d'eau aux très grands cours d'eau, sur toutes natures de substrats géologiques. 744 relevés ont été utilisés, couplant comptages de biofilms diatomiques et un assortiment complet de descripteurs abiotiques incluant les descripteurs physico-chimiques et chimiques de l'eau à la station. Afin d'assurer une bonne comparabilité temporelle entre échantillonnage biologique et l'intégration des variables abiotiques, la moyenne de chaque variable abiotique avait été calculée à l'époque sur un intervalle de temps de 2 mois (de 45 jours avant l'échantillonnage biologique à 15 jours après cette date).

3.3.1. IBD et PO₄ :

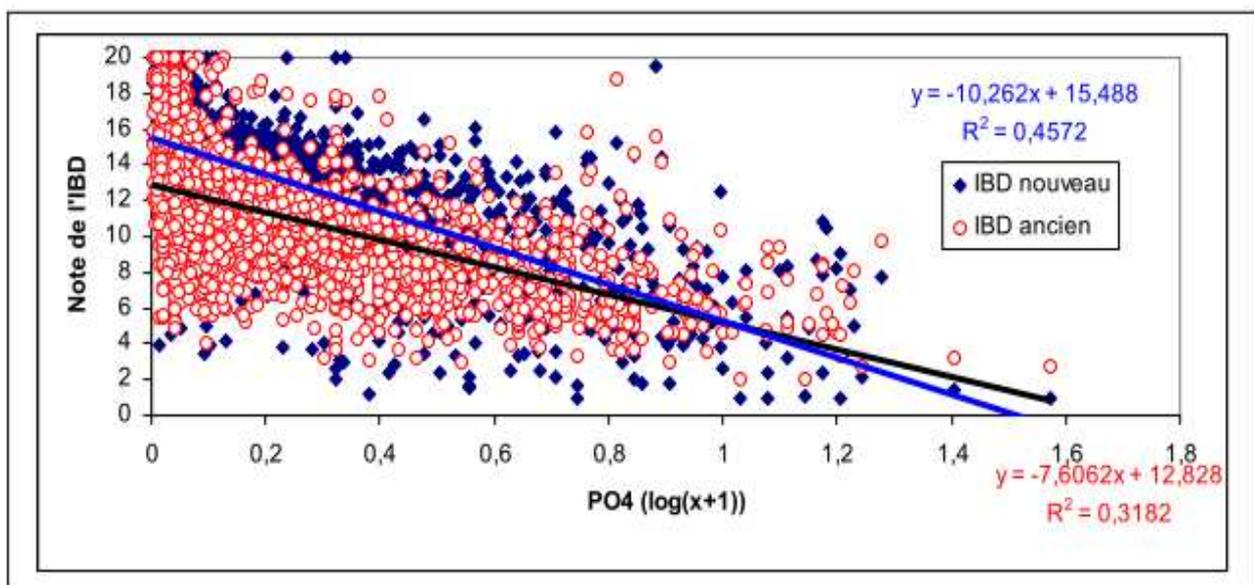


Figure 11: Relation pression-impact entre les 2 versions successives d'IBD et le log (PO₄+1)

Sur cette figure, bâtie sur la base d'une transformation logarithmique des données de concentration en PO_4 , il est possible de constater que l'IBD₂₀₀₇ (en bleu) montre une pente plus importante et un coefficient de détermination assez nettement amélioré par rapport à l'IBD₂₀₀₀ (points rouges et ajustement linéaire figuré en noir) ; cette progression venant là aussi étayer, au niveau national, le remplacement de la version précédente par le nouvel IBD₂₀₀₇.

3.3.2. IBD et NH₄ :

En utilisant exactement le même jeu de données (744 échantillons), nous avons illustré les relations entre les 2 versions d'IBD et la concentration de NH₄ (transformation logarithmique), tel que présenté dans la **Figure 12** ci-dessous :

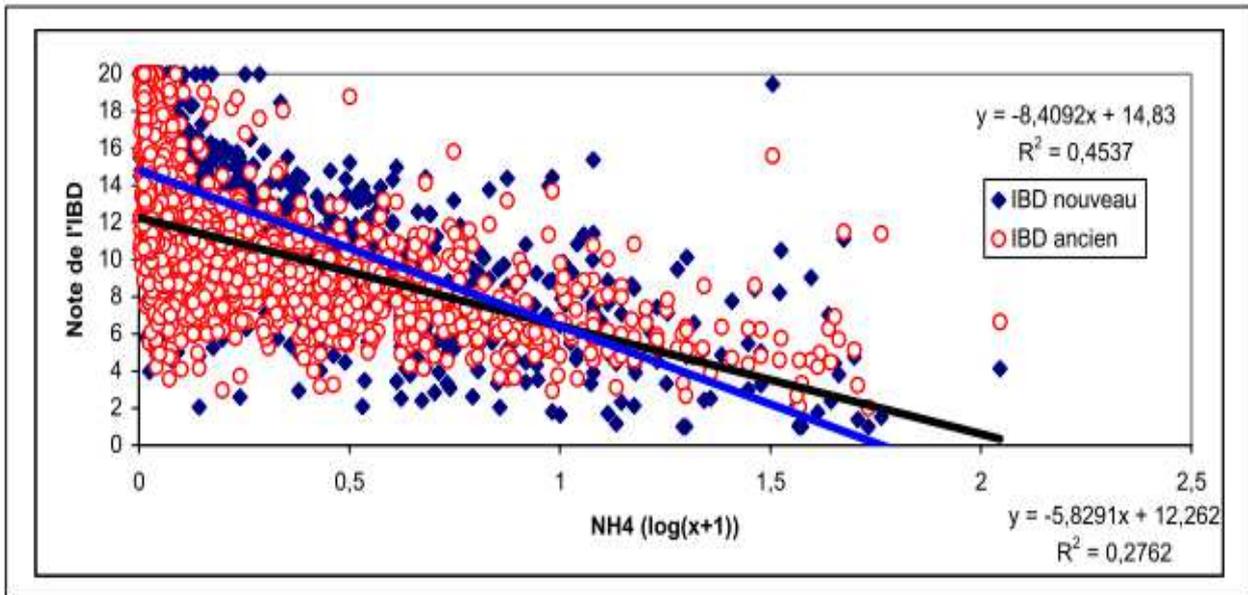


Figure 12: Pressure-impact relationships between the 2 successive versions of IBD and $\log(NH_4 + 1)$

Nous pouvons rappeler quelques limitations qui sous-tendent ces résultats, déjà soulignées auparavant :

- NH_4 agit rarement seul sur les communautés biologiques, mais de façon cumulée avec d'autres facteurs de stress anthropiques,
- Le résultat d'un ou deux échantillonnages ponctuels de l'eau est assez peu intégratif de la dimension temporelle, fait aggravé par la variabilité temporelle plus forte pour ce nutriment que pour d'autres.

Même si plusieurs valeurs outliers viennent sensiblement altérer la qualité de cette relation, ce qui est logique compte-tenu des problèmes de représentativité temporelle liée aux analyses ponctuelles de chimie, la version précédente d'IBD (l'IBD₂₀₀₀) montre une relation claire avec des concentrations NH_4 (log-transformées).

Sans aucune élimination d'outlier, l'IBD₂₀₀₇ obtient de meilleurs résultats, atteignant au moins la même qualité de relation avec ce nutriment que l'IPS ($R^2 > 0,45$), considéré comme indice de référence. Ce coefficient de détermination pourrait fortement progresser avec le retrait d'un certain pourcentage d'outliers vis-à-vis de cette relation, probablement provoqués par un prélèvement de chimie faiblement représentatif des conditions intégrées à cette station sur l'épisode et/ou par l'influence dominante d'autres types d'altération anthropique (comme DBO_5 , PO_4 , etc...)

3.4. Relations pression-impact entre IBD2007 et nutriments sur TGCE

- Ce travail a été réalisé une première fois sur la base des données disponibles durant l'été 2015 (sur une base de **523 relevés complets « TGCE »**, requêtés au seuil de surface de 8 000 km² à la station), afin de répondre à la demande du GIG « Large Rivers » dans les temps requis (soit avant fin Septembre 2015). Ce sont ces relations pression-impact, établies sur des moyennes annuelles de chimie, qui ont été envoyées au GIG vers la mi- Septembre 2015, en conformité avec sa requête (cf. paragraphes 3.4.1 et 3.4.2).

- Il a fait l'objet d'une réactualisation au terme le plus tardif possible dans le cadre de l'échéancier de réalisation de l'action, permettant d'intégrer les derniers compléments de données de chimie qu'il était possible de rapatrier à partir de PANDORE en Juillet 2017 (soit un total réactualisé de **667 relevés complets** « TGCE » requêtés au même seuil de surface de 8 000 km² à la station). Sur cette base la plus complète disponible, un travail spécifique a été réalisé afin de mettre en comparaison directe l'intégration de chimie au pas de temps annuel et au pas de temps trimestriel, a priori plus convenable pour ce maillon biologique (cf. paragraphes 3.4.3 et 3.4.4).

3.4.1. IBD₂₀₀₇ et PO₄ (analyses 2015, moyenne annuelle)

Comme pour la Figure 11, les données de PO₄ strictement obtenues sur des sites-candidats au statut TGCE (seuil de taille : 8 000 km²) ont subi une transformation log avant mise en relation avec les valeurs d'IBD₂₀₀₇ (cf. **Figure 13** ci-dessous).

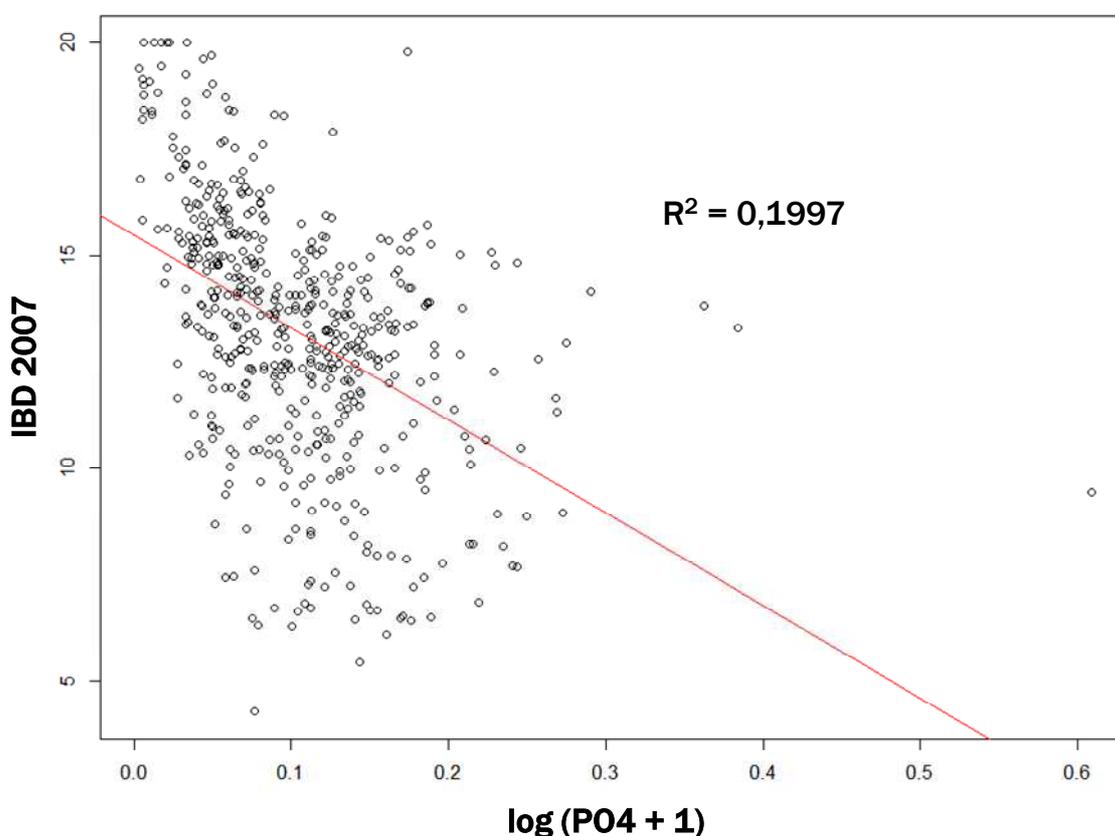


Figure 13: Relation linéaire entre IBD₂₀₀₇ et log (PO₄ + 1) (moyenne annuelle) sur le jeu de données « Très Grands Cours d'Eau »

Si une relation entre ces 2 variables ne fait pas de doute, à cette échelle de taille de cours d'eau, elle est indéniablement plus faible que la précédente illustrée en 2007 sur la base de tous les types et toutes les tailles de rivières (cf. Figure 11).

Plusieurs raisons peuvent contribuer à ce résultat:

- Comme déjà indiqué, par construction, la réponse de l'IBD₂₀₀₇ est composite et intégrative de 7 variables de pression différentes. Cet indice présentant une sensibilité spécifique différenciée vis-à-vis de chacun de ces 7 descripteurs, la réponse au seul PO₄ ne peut pas être complètement représentative du niveau réel global de la pression anthropique intégrée dans ces 523 cas d'étude.
- Comme cela peut être constaté à travers la réduction de la gamme couverte par l'échelle de log (PO₄ + 1) dans les 2 graphiques comparables (cf. Figure 11 et la Figure 13 : passage de 1,6 à 0,55 dans la même échelle log, soit une réduction d'environ 2/3 du gradient national couvert dans cette échelle), la sélection de sites uniquement sur Très Grands Cours d'Eau, même dans une acception élargie (8 000 km² place du seuil de 10 000 km² défini au niveau européen pour l'exercice d'intercalibration en cours) réduit fortement le gradient de concentration en PO₄ couvert.

- Par ailleurs, un échantillon de diatomées benthiques serait *a priori* plus influencé par une qualité intégrée de l'eau écoulee pendant 2 à 3 mois au maximum que par une moyenne annuelle. Ce point sera vérifié dans la suite des travaux, mais se référer à une moyenne annuelle s'éloigne probablement plus de la qualité de l'eau intégrée à laquelle les diatomées ont été réellement soumises au cours du cycle de croissance d'un biofilm estival.

Compte tenu de la première limitation évoquée ci-dessus, nous avons essayé de procéder à une suppression des points les plus éloignés de la relation PO_4 -IBD présentée Figure 13, afin d'étudier la progression de la qualité de la relation avec la suppression des points considérés comme les outliers les plus manifestes (ce n'est pas anormal qu'il y en ait, notamment puisque le PO_4 n'est que l'un des 7 descripteurs de pression anthropique utilisés par construction pour caler la réponse de l'IBD).

Sur la base du retrait automatique sous logiciel R des cinq points jugés les plus aberrants parmi les 521 valeurs prises en compte sur ce paramètre (points repérés comme les 1% d'outliers principaux), l'évolution de la relation s'effectue comme illustré en **Figure 14** :

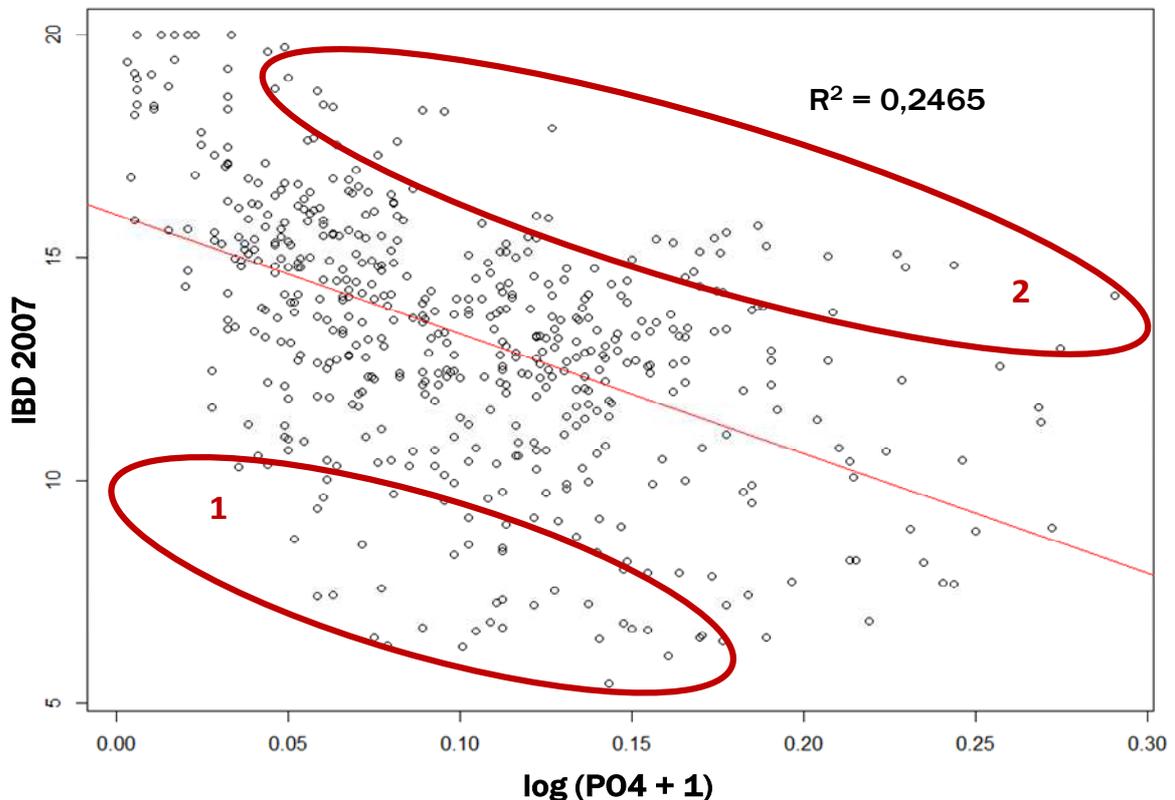


Figure 14: Relation linéaire entre IBD_{2007} et $\log(PO_4 + 1)$ (moyenne annuelle) en Très Grands Cours d'Eau après enlèvement d'1% d'outliers

Dans le cas de figure représenté, le principe de sélection automatique par l'appli sous R des valeurs jugées les plus aberrantes pose question, dans la mesure où l'enlèvement de 3 des 5 données jugées « outliers » suffit pour réduire par 2 le gradient des valeurs de PO_4 représentées, supprimant d'ailleurs les plus fortes valeurs rencontrées alors qu'elles doivent être importantes pour établir la relation réelle qui s'établit sur la note d'IBD sous influence principale de la concentration en PO_4 (en effet, c'est probablement dans le secteur du toit du nuage de points -cf. ellipse 2- que les réponses IBD sont les moins bruitées par une réponse multi-paramètres).

Or le but d'un indice ou d'une méthode nationale d'évaluation est aussi d'arriver à évaluer correctement des milieux aquatiques subissant une forte altération anthropique (ici, les plus forts niveaux de PO_4 rencontrés dans tout le jeu de données). Dans le cadre de cette étude nationale TGCE, cette réduction importante du gradient de concentrations en PO_4 pris en compte, qui plus est en base transformée log, n'est pas vraiment justifiable sur un plan scientifique ou technique, et constitue une gêne vis-à-vis de l'étude de relation pression-impact. En effet, sur ces cours d'eau de grande taille et avant le retrait de ces valeurs outliers, on rencontre déjà, par rapport à des cours d'eau plus petits, une absence quasi-totale de situations de référence vraie et une diminution des situations très fortement dégradées, du fait des lois de dilution des rejets et de l'effet plus perceptible des processus d'auto-épuration sur des hydrosystèmes plus longs à temps de résidence accrus.

De plus, à l'observation comparée des Figures 13 et 14, ces 3 points retirés, qui contribuent à diminuer le gradient, ne sont pas forcément de vraies valeurs aberrantes vis-à-vis de la relation, réelle ou supposée, entre PO₄ et IBD₂₀₀₇. En effet, ces points respectent la pente de la relation moyennant un décalage vers un plus faible impact biologique, et leur statut d'outlier pourrait en fait être représentatif de situations à effet monofactoriel ou quasi-monofactoriel des orthophosphates à fortes concentrations, par rapport à beaucoup d'autres points multi-impactés qui entrent dans le calcul de cette relation. L'impact multifactoriel est en effet le plus souvent la norme à l'aval de STEPs, qui rejettent en mélange différents paramètres d'enrichissement des eaux en nutriments et peuvent aussi provoquer un déficit en O₂, la note IBD s'établissant sous influence de ce faisceau multiple de pressions.

Quoi qu'il en soit, si le coefficient de détermination de la relation a progressé de 4,5 % avec la suppression des 5 points répertoriés comme principaux outliers, cette progression est plutôt modeste et au final, la qualité de la relation peut être qualifiée de faible à modérée sur notre assortiment national de TGCE. Les raisons explicatives les plus plausibles, déjà évoquées plus haut, sont rappelées ci-dessous :

- La raison probablement principale est qu'un seul descripteur de pression anthropique (PO₄) est pris en compte pour l'établissement de cette relation, alors que la réponse biologique de l'IBD₂₀₀₇ a été initialement calée à partir de l'effet intégré de 7 descripteurs différents ;
- Une autre hypothèse susceptible de diminuer la qualité de la relation est l'utilisation de la moyenne annuelle de PO₄ au lieu d'une intégration sur un temps plus court (2 à 3 mois), plus en phase avec l'écologie des biofilms prélevés.

Dans le cadre de l'exercice européen d'intercalibration 2015-2017 qui nous a concernés, c'est l'unique gradient de concentration en PO₄ qui a été utilisé pour intercaler les méthodes nationales d'évaluation des très grands cours d'eau. Outre le fait que sur TGCE, à l'échelle européenne comme à l'échelle Française, on assiste à une nette compression de la représentation de ce gradient de nutriment par rapport aux hydrosystèmes plus petits, il est évident aussi que les réponses des outils d'évaluation conçus pour évaluer des gradients multiples d'anthropisation peuvent fournir des réponses éventuellement décalées et probablement plus floutées que d'autres systèmes ou métriques uniquement dédiées à la mesure d'eutrophisation.

Dans ce contexte particulier d'exercice inter-Etats-Membres à enjeu de vérification et d'ajustement de niveau de calage de systèmes nationaux d'évaluation, qui constitue sujet un peu sensible, il nous a semblé utile de pousser un peu plus les réflexions sur la réponse de notre outil indiciel national à cette pression particulière dans un contexte multipressions, afin de mieux comprendre et assimiler son comportement sur les hydrosystèmes de cette taille et d'être mieux en mesure d'interpréter les significations possibles du positionnement de points se situant dans des secteurs particuliers des biplots présentés (Figure 13 et Figure 14), voire d'analyses collectives faites sur le jeu de données d'intercalibration dans le cadre de l'exercice.

Par ailleurs, suite à l'effet inattendu qu'a revêtu la suppression de seulement 5 outliers déterminés sous R (3 valeurs enlevées ayant suffi pour réduire d'un facteur 2 le gradient de valeurs en échelle log décimal, soit une compression encore sensiblement plus importante en valeurs réelles de concentrations), un questionnement méthodologique manifeste se pose sur la façon la plus adéquate de supprimer les valeurs aberrantes pour conforter la réponse pression-impact dans de telles analyses. En effet, la question « qu'est-ce qu'une valeur aberrante dans un tel biplot ? » renvoie immédiatement derrière à « aberrante par rapport à quelle pression ? ». Sachant que, dans notre jeu de données national comme dans les jeux de données des autres Pays-Membres, même si l'on étudie une relation pression-impact en monofactoriel (ici, le PO₄), il est certain que la notation indicielle obtenue sur un bon nombre de points subit une altération sur un plan plus composite, comme cela se passe très classiquement dans différentes configurations rencontrées dans les réseaux de surveillance de pas mal de pays-membres, et notamment à l'aval de STEPs.

Dans ce sens, la partie supérieure du nuage et sa pente sont probablement sous influence dominante de situations de monopression PO₄, et le cumul de pressions contribue probablement à obtenir des notations IBD multi-impactées, donc plus basses. Il semble logique de présumer que, si la baisse d'IBD était vraiment et uniquement due à l'augmentation des teneurs de PO₄, la régression linéaire devrait couper l'axe Y dans un secteur proche de la note d'IBD de 20 et les notes d'IBD devraient baisser, avec un certain flou statistique dû au déficit de représentativité temporelle des échantillonnages ponctuels de chimie, en fonction de l'augmentation des teneurs en PO₄. Les points les plus anormaux vis-à-vis de cette relation ne sont donc probablement pas les 3 points parmi 5 qui ont été retirés entre la Figure 13 et la Figure 14, qui réduisent fortement le gradient de PO₄ représenté et modifient très certainement le niveau d'ajustement de cette relation cause-effet, mais plus probablement des points situés dans le secteur de l'ellipse N°1 (note d'IBD très déprimée pour de relativement faibles concentrations de PO₄, qui ne peuvent donc pas à elles seules être désignées comme responsables du faible niveau obtenu par la note d'indice diatomique dans de tels cas).

L'explication de valeurs anormalement basses d'IBD, tout particulièrement dans le secteur de l'ellipse 1, tient probablement à des **valeurs élevées de pression anthropique composite** mettant aussi en jeu d'autres descripteurs hydrochimiques sous-jacents, dont certains sont connus pour être aussi impactants ou presque aussi impactants que PO₄ pour les diatomées, comme la DBO₅, le NH₄, l'azote Kjeldahl, le déficit de saturation en oxygène *in situ* (% Sat O₂), **et/ou** à une **mauvaise représentativité temporelle des prélèvements de chimie** qui conduiraient, malgré l'intervention de moyennes, à sous-estimer fortement les concentrations intégrées en PO₄ qu'auraient réellement subi les biofilms prélevés *in situ*.

Cette bonne représentativité temporelle des échantillons ponctuels d'eau n'étant pas garantie, *a priori*, il n'est pas non plus à exclure qu'elle provoque quelques positionnements outliers dans la zone du toit du nuage (secteur de l'ellipse 2). Il est en effet envisageable que momentanément, au moment du prélèvement et pour cause d'épisodes hydrologiques plus ou moins marqués accélérant les transferts (épisodes de crue), la teneur en orthophosphates de la colonne d'eau restituée grâce aux échantillonnages ponctuels prélevés puisse donner **un résultat de calcul statistiquement surévalué** par rapport à la **réalité de l'exposition moyenne** aux orthophosphates *in situ*, intégrée sur la période de croissance du biofilm correspondant, conduisant à un positionnement de quelques relevés à un niveau exceptionnellement modéré de réponse biologique par rapport à la moyenne calculée de PO₄.

Mais ce cas de figure, qu'on aura le plus de chances de repérer sur le toit du nuage (secteur de l'ellipse 2), est finalement assez improbable à l'échelle du jeu de données : 1) vu la pratique de la moyenne sur les données disponibles, qui vient pondérer la représentation d'une valeur trop forte, et 2) vu la très nette dominance, dans un hydrogramme annuel, des moments de débits faibles à moyens par rapport aux pics de concentration de nutriments.

En effet, ces événements très transitoires de pics de concentration en nutriments interviennent souvent lors de la courte phase ascendante d'un pic de crue, lorsque les transferts liés à une lixiviation rapide dominent encore sur les phénomènes de dilution provoqués par l'onde de crue. Il représentent donc très peu de temps sur un hydrogramme annuel et ont statistiquement moins de chances d'être prélevés que d'autres configurations hydrochimiques en conditions plus stabilisées.

D'ailleurs, à l'appui de cette considération sur le faible risque de rencontrer des cas de figure de **surestimation de l'exposition en nutriments**, sur la base des **résultats analytiques intégrés par rapport à l'exposition réelle *in situ***, l'examen de la Figure 13 (avant tout enlèvement d'outliers) n'en fait apparaître qu'un sur les 5 principaux outliers repérés (le point isolé en haut de graphe avec la note d'IBD très proche de 20) : la teneur moyenne en PO₄ quantifiée dans les échantillons ponctuels, ici transformée log, étant probablement sur-représentée par rapport à la concentration intégrée réellement subie par le biofilm lors de son épisode de croissance *in situ*.

Ce premier examen des relations pressions-impacts sur les TGCE de notre territoire national s'est opéré en contexte de temps très contraint, puisque le maximum d'information contextuelle possible était requis en tout début d'exercice d'intercalibration UE (en même temps que la fourniture des jeux de données nationaux) afin de venir appuyer la qualité d'évaluation permise par les systèmes nationaux d'évaluation soumis à l'exercice et en vue de démarrer rapidement celui-ci. C'est principalement dans ce cadre qu'il pouvait être intéressant de mettre en œuvre une stratégie adaptée d'enlèvement d'outliers, afin de conforter la relation pression-impact relative à notre outil national.

Cependant, si les résultats des relations pression-impact ont fait l'objet d'une actualisation ultérieure, sur la base de l'accroissement des doublets de données disponibles, il n'a pas été possible de reprendre un travail de fond visant à éliminer les valeurs aberrantes selon une méthode biomathématique plus convenable vis-à-vis de la nature des relations étudiées. En effet, la relative faiblesse des relations pression-impact illustrées, à notre niveau national et plus encore au niveau du GIG, a 2 origines principales :

- **la compression du gradient de pression** rencontré sur ces gros hydrosystèmes,
- **l'aspect très souvent composite des pressions anthropiques** qui s'exercent au même site.

Donc, au bilan, le niveau d'expression de relations pression-impact sur TGCE, ainsi que leur amélioration éventuelle, tiennent moins à une évolution de la stratégie d'enlèvement d'outliers qu'aux 2 types de facteurs structurels précités. Bien qu'existante, la marge de progression permise par cet enlèvement plus judicieux d'outliers implique une forte intervention d'avis-expert et reste probablement modeste. Dans un contexte général de tension sur la ressource humaine et les échéanciers de travail, vu le niveau de gain envisageable, probablement modeste, et le faible enjeu une fois l'exercice d'intercalibration réalisé, cette piste pour l'amélioration de relations « pression-impact » n'a donc pas été poussée plus avant.

Au cas où cette orientation de travail serait reprise à l'avenir, les orientations suivantes pourraient être suivies pour conduire à des résultats un peu améliorés :

- Afin d'améliorer la régression de l'IBD avec le seul descripteur PO_4 , selon la même logique que ce qui a été fait dans l'exercice d'intercalibration, Il faudrait probablement concentrer l'enlèvement d'outliers, non pas en fonction du premier ajustement linéaire produit sur la base de toutes les données brutes, mais en sélectionnant et en éliminant en premier lieu les points ou les réponses de l'IBD sont manifestement provoquées de façon dominante par d'autres descripteurs que par les orthophosphates (vérifications à faire en amont, par exemple par seuillage sur les autres descripteurs de pression anthropique).
- Si l'optique est d'optimiser, sur TGCE, la relation entre l'IBD et un contexte multi-pression mettant en jeu tout ou partie des 7 descripteurs d'anthropisation utilisés pour sa genèse, une voie d'approfondissement pourrait être de s'appuyer sur l'étude de régression multifactorielle (par exemple, piste méthodologique des régressions PLS), qui permettent de prendre en compte et de tenter d'isoler l'influence des différents facteurs contributifs. Néanmoins, si cette technique peut apparaître très séduisante sur le papier pour résoudre le problème rencontré, 1) elle restera confrontée aux défauts de représentativité temporelle présentés par les analyses chimiques de l'eau ; 2) elle sera de toute façon impactée par les co-occurrences factuelles de facteurs regroupés que l'on trouve très souvent en mélange composite variable juste à l'aval de STEPs ; et 3) l'influence relative accordé à chaque paramètre à un site donné en fonction des résultats d'analyses chimiques locales sera la résultante intégrée pour ce paramètre obtenue au niveau du jeu de données total, et non le poids qu'exerce réellement ce paramètre au site précis d'observation que l'on cherche à évaluer. Ainsi, il sera utilisé une grille d'affectation commune de poids qui ne pourra pas être du même niveau de représentativité et de légitimité selon des typologies de situations résultant : 1) de pollutions diffuses agricoles ou domestiques mises en place à l'échelle de larges territoires d'une part, et 2) de rejets ponctuels de STEPs ou agroindustriels d'autre part.
- Un autre type d'approche plus similaire au principe de genèse de l'IBD pourrait passer par l'utilisation d'une métrique composite d'anthropisation rendant compte de façon intégrée de l'influence combinée de plusieurs descripteurs de pression anthropique sur le jeu de données TGCE (dont les facteurs de forçage ne s'exercent probablement pas tout-à-fait dans le même équilibre que sur de plus petits cours d'eau).

3.4.2. IBD₂₀₀₇ et NH_4 (analyses 2015, moyenne annuelle)

Sur ce même jeu de données disponible à l'été 2015, uniquement constitué de relevés TGCE, nous avons également essayé de représenter la relation obtenue entre IBD₂₀₀₇ et concentrations de NH_4 (valeurs log-transformées). Cette relation est présentée en **Figure 15**.

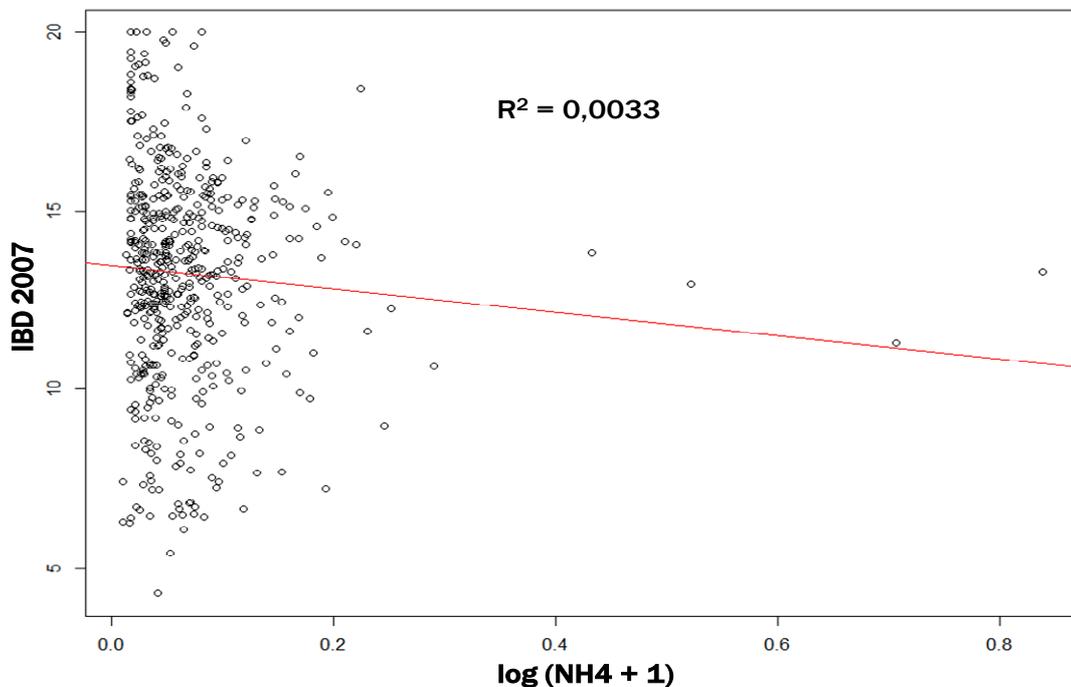


Figure 15: Relation linéaire entre les notes d'IBD 2007 et le $\log (NH_4 + 1)$ (moyenne annuelle) dans les Très Grands Cours d'Eau

Si à l'échelle des cours d'eau de France, toutes échelles de taille confondues, cette relation est indéniable (cf. Figure 12), à l'échelle des TGCE étudiée ici, l'influence de ce paramètre (et/ou la représentativité spatio-temporelle des valeurs de ce paramètre obtenues dans les analyses) semble nettement atténuée et bruitée. Cependant, si le R^2 très faible est révélateur d'un niveau important de résidus par rapport à l'ajustement linéaire modélisé (qui n'est donc pas convenable), une pente lie malgré tout l'accroissement des teneurs en NH_4 et la diminution des notes d'IBD.

En prenant du recul par rapport à l'ajustement linéaire modélisé, si l'existence d'une relation est bien suggérée via la décroissance des notes d'indice avec les plus fortes valeurs de NH_4 (partie haute du nuage de points) :

- elle semble très modeste sans un retrait des valeurs d'IBD très certainement liées à d'autres paramètres d'altération anthropique que le NH_4 (secteur des plus basses valeurs indicielles). En effet, il semble clair que d'autres types de nutriments provoquent ici de plus fortes baisses d'indice par rapport au NH_4 et génèrent du bruit,
- le coefficient de détermination, très faible, est non satisfaisant.

Trois hypothèses relatives à cet état de fait peuvent être évoquées :

- La moyenne annuelle de NH_4 pourrait ne pas être assez représentative de l'exposition réelle à la saison précise où les biofilms des échantillons ont été exposés,
- Dans les grandes rivières, en général, il devient plus difficile de trouver des valeurs élevées à très élevées de NH_4 du fait de la réglementation sur les rejets de sortie de stations d'épuration, qui conditionnent depuis 30 ans des améliorations subventionnées des infrastructures et procédés épuratoires ; de la forte dilution des sorties d'égouts dans de tels grands systèmes ; des processus de minéralisation et de nitrification *in situ* qui interviennent lors des processus d'auto-épuration, qui conduisent à des concentrations plus modérées dans ces hydrosystèmes en liaison avec les temps de résidence. Le gradient est alors moins étendu que dans les petites et moyennes rivières.
- Comme déjà indiqué précédemment, dans un secteur comparable à celui de l'ellipse N°1 dans la Figure 13, des valeurs très basses d'indice ne peuvent pas être occasionnées par les valeurs faibles de NH_4 analysées dans les mêmes eaux, et sont donc probablement dues à d'autres paramètres de pression. Une élimination sélective de valeurs aberrantes dans ce secteur précis conforterait très certainement la pente du modèle et le coefficient de détermination.

Quoi qu'il en soit, ces relations sont beaucoup plus pauvres que sur un jeu de données national basé sur tous les types de cours d'eau (voir les Figures 11 et 12).

Comme cela a été évoqué pour le PO_4 , il est possible que l'utilisation de valeurs moyennes trimestrielles, mieux calée sur la durée d'exposition d'un biofilm en fonction de sa durée de vie sur son substrat, améliore dans une certaine mesure les relations. Cette façon de faire différente et probablement plus adaptée, a donc été testée au chapitre suivant.

D'autres pistes qu'il serait possible de travailler passent par la suppression de certaines valeurs faibles du nutriment, sous condition que les valeurs fortes rencontrées pour d'autres descripteurs anthropiques permettent de justifier qu'elles sont responsables de la baisse de note indicielle constatée au relevé.

Comme déjà indiqué, on pourrait aussi essayer de tester sur TGCE la relation entre l'IBD₂₀₀₇ et un gradient de pression composite bâti avec le même assortiment et le même poids relatif des variables anthropiques qui ont été utilisées pour son développement. Enfin, l'utilisation de régressions quantiles pourrait aussi contribuer à une meilleure démonstration de la relation pression impact.

3.4.3. IBD₂₀₀₇ et nutriments (analyses 2017, comparaison des temps d'intégration)

Suite aux compléments de requêtes réalisés en Juillet 2017, il a été possible d'augmenter l'assise-données totale théoriquement disponible sur les TGCE nationaux (soit 667 relevés fait sur des sites de plus de 8 000 km^2 , dont 550 sur sites de plus de 10 000 km^2) avec données de chimie existantes. Cependant, différents problèmes rencontrés sur les données abiotiques ont nécessité l'adoption de pratiques visant à conforter leur fiabilité, conduisant à une réduction d'assise effective des données conservées pour les analyses.

Un premier post-traitement a été fait à partir des codes-remarques présentant des anomalies mettant en doute la fiabilité du contenu de la colonne de valeur numérique d'un nutriment donné, à savoir l'existence d'une valeur numérique assortie du Code 0 (mesure non faite) ou du Code 3 (mesure de saturation) improprement affecté à un autre descripteur que la saturation d' O_2 , code-remarque ininterprétable pour d'autres paramètres et mettant en doute la fiabilité du contenu de la valeur numérique associée.

Globalement, ces 2 informations font apparaître une incohérence entre la colonne porteuse d'une information numérique sur une analyse et le code-remarque associé. Malgré que la valeur analytique semble la plupart du temps quantifiée et du domaine du plausible, l'incohérence du code-remarque fait planer un doute sérieux sur la chaîne de stockage et rapatriement des données (contenu de la base initiale ? problème à la requête ?). Or, il ne nous est plus possible de tracer le pourquoi du problème ni son historique une fois ce contenu intégré dans la BDD PANDORE, dont nous ne sommes pas les gestionnaires. Dans un tel cas, la valeur numérique correspondante n'a pas été utilisée, faisant perdre le bénéfice d'une partie non-négligeable des relevés.

D'autre part, la colonne « Code-remarque » faisait souvent apparaître un code-remarque 2 (< Seuil de Détection ou SD), ou un code-remarque 10 (< Seuil de Quantification ou SQ). La valeur correspondante dans la colonne des valeurs numériques n'est donc pas le résultat d'une analyse, mais l'indication d'un seuil de détection ou de quantification analytique non atteint.

Au moins dans le cas des analyses de nutriments, il ne peut pas être donné d'interprétation quantitative différente de ces 2 notions, qui ont donc été considérées ici comme équivalentes. Mais, afin de ne pas avoir un nombre très conséquent de valeurs numériques manquantes (valeurs analytiques non-quantifiées), ce qui constitue ensuite un obstacle sérieux pour la réalisation de diverses catégories d'analyses, il a fallu malgré tout affecter une valeur numérique à ces cellules lorsque les informations semblaient cohérentes et normales.

Il a tout d'abord été procédé à un enlèvement des valeurs SD ou SQ anormalement élevées (laboratoires réalisant les analyses avec une sensibilité éventuellement adaptée à d'autres natures de matrices, de type eaux usées, boues, sédiments, mais **manifestement insuffisante et non-convenable pour une matrice d'eaux naturelles**). Les plafonds de SD ou SQ tolérés ont été de : 0,5 mg/l pour NH₄ ; 0,2 mg/l pour PO₄ ; 3 mg/l pour NO₃. Toutes les valeurs de SD ou SQ plus élevées que ce qui précède, révélatrices d'une sensibilité de méthode complètement inadaptée au contexte des analyses d'eaux naturelles, ont donc été éliminées.

Les autres valeurs de seuils analytiques non atteints (SD, SQ) et restant en-deçà des valeurs-plafond précitées ont été divisées par 2, en vue de combler la lacune en valeur numérique vraie et de tenter de préserver autant que possible (avec cette pratique, il reste malgré tout des chevauchements...) un gradient statistique quantitatif entre valeurs non quantifiées et valeurs numériques réellement analysées.

Il a ensuite été étudié l'incidence de différents temps d'intégration abiotique pour générer des relations pressions-impact (cf. Tableau 2), tant en termes d'assise de chimie prise en compte qu'en termes d'évolution qualitative des relations pression-impact étudiées.

En vue de comparaison directe avec des éléments antérieurs obtenus sur un jeu de données national intégrant les cours d'eau de toutes tailles et de tous types, déjà présentés précédemment dans le même rapport (cf. Figure 11 et Figure 12), les analyses de relations présentées en **Figure 16** (page suivante) ont été effectuées à titre d'exemple vis-à-vis de 2 nutriments disponibles dans la grande majorité des analyses, à savoir le PO₄ et le NH₄.

Il a tout d'abord été étudié l'assise statistique des données abiotiques intégrées pouvant être mises en face des relevés biologiques disponibles en fonction du créneau d'intégration temporelle retenu (cf. **Tableau 2**).

Tableau 2 : Nombre moyen de valeurs de chimie impliquées dans le calcul des valeurs moyennes

Durée d'intégration temporelle	PO ₄	NH ₄
annuelle [j-365]-[j+15]	9.53	9.53
4 mois [j-105]-[j+15]	3.79	3.78
3 mois [j-75]-[j+15]	2.9	2.9
2 mois [j-45]-[j+15]	2.0	2.0

Pour une surveillance classique de sites de réseaux, de 10 à 12 analyses de physico-chimie et de chimie sont habituellement réalisées annuellement.

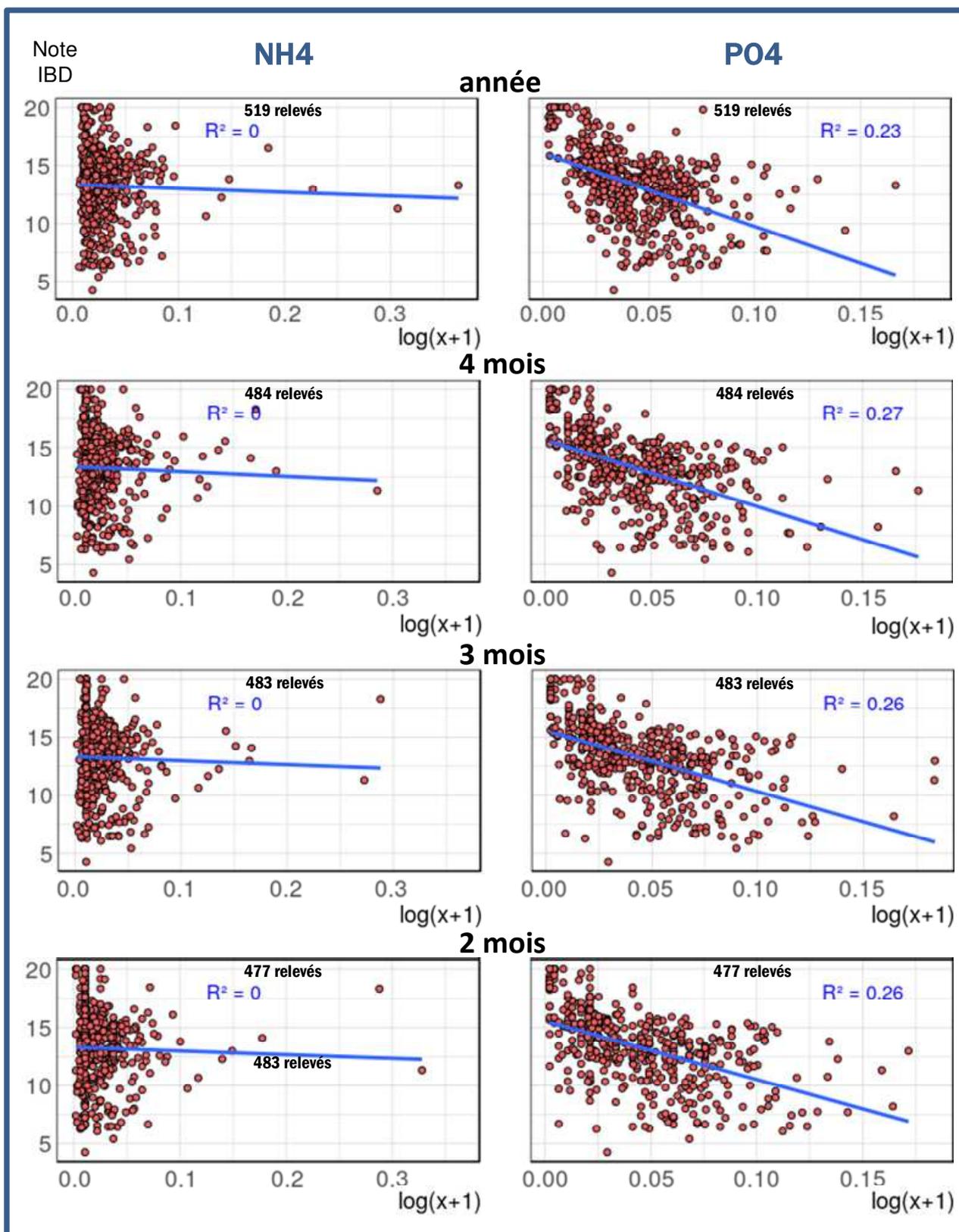


Figure 16 : Relations pression-impact établies entre IBD et nutriments sur TGCE de France (requête PANDORE Juillet 2017)

En pratique, nous avons établi la requête sur la **moyenne au pas de temps annuel** selon l'application de la fenêtre temporelle indiquée en haut de Figure 1 (J- 365 – J+15), soit un calcul de moyenne sur une durée de 380 jours. Les diminutions de durée successivement prises en compte pour tenter de mieux correspondre à la durée d'intégration biologique du maillon diatomique figurent dans la colonne la plus à gauche du **Tableau 2**, soit 4 mois (120 jours), 3 mois (90 jours) et enfin, 2 mois (60 jours).

Malgré l'allongement du créneau temporel à une quinzaine de jours juste après le prélèvement biologique, donc sous la même influence saisonnière et le plus souvent aussi sous la même tendance de régime dans le cours d'eau que le prélèvement biologique réalisé (période de belle saison et de moyennes à basses-eaux), le nombre d'analyses pour l'établissement d'une moyenne annuelle s'établit légèrement en-dessous de la pratique nationale de surveillance physico-chimique de référence, du fait de l'intervention de quelques sites du réseau de référence R-REF (année 2007 + un reliquat de suivi en 2008), qui faisaient l'objet de 4 analyses par an seulement, et de quelques cas où cette requête annuelle faite précisément de date à date bien a pu laisser échapper certaines analyses à quelques jours près.

On peut constater que les 2 créneaux intermédiaires étudiés ici donnent une assise intéressante pour calculer une moyenne présentant un certain degré de robustesse (3,8 analyses disponibles pour 4 mois et 2,9 analyses pour 3 mois), dans un créneau d'intégration temporelle qui se rapproche du temps de vie habituel d'un biofilm diatomique.

En raccourcissant encore plus la fenêtre temporelle, l'intégration à 2 mois semble donner accès au calcul d'une moyenne sur 2 valeurs en moyenne sur le réseau, ce qui serait une base assez correcte si cela restait la règle. En réalité, elle pose déjà un problème assez net de perte de représentativité temporelle de l'ambiance chimique pour un nombre croissant de sites. En effet, de façon sous-jacente mais bien réelle, certains sites présentant 3 valeurs sur la chronique viennent masquer des sites où la représentativité temporelle des analyses chimiques devient plutôt aléatoire (effectif croissant de sites ne disposant plus que d'une seule valeur de chimie, parmi lesquels certains sites de référence, prélevés nettement moins souvent que les autres, mais aussi d'autres sites de réseaux classiques...). L'affaiblissement de l'assise chimie occasionné par le raccourcissement du créneau temporel se matérialise encore plus concrètement par la mise en place d'une perte complète de 6 relevés diatomiques faute de toute date de prélèvement chimique entrant dans la fenêtre temporelle ciblée (cf. bas de la Figure 16 : total diminué à 477 relevés disponibles). Sur le plan biologique d'autre part, si 45 à 60 jours constituent une durée permettant l'installation d'un biofilm développé en 3D et déjà relativement mature, la durée d'un cycle de croissance complet d'un biofilm donné sur un substrat dur, avant bullage et décollement, peut atteindre 3 à 4 mois, soit une capacité intégrative d'environ le double de cette durée. Dans notre cas, finalement, l'aspect jugé le plus problématique pour l'analyse calée sur ce pas de temps relativement court consistait en l'appauvrissement et la perte de robustesse de l'intégration chimique réalisée.

C'est en fonction de ces éléments de cadrage des agrégations de données physico-chimiques réalisées que les ajustements linéaires obtenus sur TGCE vis-à-vis de 2 paramètres d'impact anthropique pris en exemple (NH₄ et PO₄), généralement assez bien corrélés aux notes d'IBD₂₀₀₇ sur des jeux de données issus de cours d'eau de plus petite taille (cf. Figures 11 et 12), ont été présentés en page précédente (cf. Figure 16).

Au travers de cette étude de l'influence de la durée d'intégration temporelle sur le niveau de réponse biologique, il est possible de percevoir la répercussion de l'allongement ou du raccourcissement du créneau pris en compte sur la disponibilité de données valides. En effet, au pas de temps annuel, 519 doublets de données sont disponibles. En procédant à un raccourcissement du créneau pour arriver plus en rapport avec le temps d'intégration biologique caractéristique d'un biofilm, on assiste à une perte sèche de doublets d'un niveau très comparable pour 120 jours (484) et 90 jours (483) d'intégration temporelle. Le raccourcissement à 60 jours fait passer un nouveau palier qui conduit à une nouvelle perte de 6 relevés (476), mais affecte surtout l'assise en analyses chimiques disponibles, qui bascule à une seule valeur sur une part non-marginale des relevés et conduit à une dégradation de l'assise permettant le calcul de valeurs moyennes robustes.

Sur le plan des relations mises en évidence, même si ces 2 nutriments jouent indéniablement un rôle structurant sur une diminution des notes d'IBD₂₀₀₇ en fonction de l'augmentation de leur concentration (on peut en particulier le vérifier en fonction de la physionomie du toit du nuage de points), ces résultats illustrent bien que sur TGCE, ces relations sont plus bruitées et de qualité moindre qu'à l'échelle de l'ensemble des hydrosystèmes de France, toutes tailles et tous types d'hydrosystèmes confondus, comme présenté en Figures 11 et 12). Si la relation avec le PO₄ conserve une certaine consistance et ne peut pas être mise en doute, celle avec le NH₄ devient nettement plus ténue du fait de la diminution importante du R² qui s'approche de 0.

Les raisons générales relatives à la dégradation de ces relations dans ce type de grands hydrosystèmes, dont certaines ont déjà été évoquées antérieurement, peuvent tenir aux raisons suivantes :

- réduction, à leurs 2 extrémités, des **gradients** des descripteurs d'anthropisation pris en compte : 1) du fait de l'absence de situations de référence vraie et du faible effectif de relevés représentatifs de références par défaut ; mais aussi 2) du fait de l'importance plus grande du facteur de dilution des rejets, de l'augmentation des temps de résidence des écoulements et des processus d'autoépuration au long de leur linéaire, réduisant les cas où l'on rencontre de fortes concentrations.

- **plus grande fréquence de situations multi-impactées sur de tels grands hydrosystèmes** collectant souvent des rejets pluviaux péri-urbains et des rejets de grosses STEPs. Ainsi, en liaison avec la stratégie nationale dominante de traitement collectif des effluents adoptée sur notre territoire, beaucoup de cours d'eau Français (dont les TGCE) sont concernés par un contexte multi-altérations qui floute la réponse que l'on peut obtenir avec cet indice vis-à-vis de la prise en compte d'un seul paramètre révélateur de pression. La réponse de l'IBD₂₀₀₇ a été initialement calée vis-à-vis d'un gradient composite de 7 descripteurs abiotiques impactés par diverses altérations anthropiques. Ainsi, l'étude de la relation exposition-réponse à un seul descripteur monofactoriel (ici, le PO₄ seul, ou le NH₄ seul...) va inévitablement engendrer des réponses biologiques parasites et des résidus indésirables vis-à-vis de la relation étudiée, à chaque fois que le site est multi-impacté (ex : aval immédiat de rejet de STEP, sous influence conjointe de plusieurs descripteurs responsables d'impacts biologiques). Ces situations conduisent à un affaiblissement de la relation monofactorielle pression-impact obtenue, dont la pente et l'ajustement point par point des relevés contributifs sont influencés et bruités par la réponse biologique à divers paramètres hydrochimiques non considérés, et dont l'ajustement de chaque relevé contributif est bruité (donc affaiblit le R²) par rapport à l'ajustement linéaire modélisé.

- par rapport à un cours d'eau plus petit, plus rapidement mélangé et homogénéisé, et en fonction de la taille de l'hydrosystème, des difficultés d'accès dans le lit et des éventuelles influences parasites de bordure, difficulté accrue à représenter correctement, à partir d'un petit nombre de prélèvements ponctuels, une exposition physico-chimique intégrée correctement représentative de la colonne d'eau sur toute la section mouillée, à un site donné faisant partie d'un réseau de surveillance.

Sur l'assise nationale de relevés TGCE, la contribution possiblement conjointe de ces 3 types de raisons conduit : 1) à un ajustement linéaire moins convaincant que pour des hydrosystèmes plus petits qui intercepte, de façon non escomptée et anormale (ajustement de pente à l'origine faussé), l'axe y vers des notes d'IBD de 15 pour le PO₄ et de 14 à 15 pour le NH₄ ; 2) à une réduction de pente des relations avec ces 2 paramètres par rapport à un jeu de données de cours d'eau toutes tailles confondues ; ainsi que 3) à la génération d'un niveau élevé de résidus qui vient dégrader le R².

Une autre information d'ordre général qui ressort est que, dans ce type d'hydrosystème, le PO₄ présente un effet plus structurant, ou moins bruité par les différents interférents précités, que l'ammonium (dont la pente est plus faible et le R² devient très proche de 0). Malgré tout, dans les 2 cas, on perçoit bien l'existence indéniable d'une action du nutriment sur le niveau de notation de l'IBD, qui se traduit par la physionomie du haut du nuage de points mis en jeu dans l'établissement de la relation. Mais on perçoit aussi le fort bruitage multiparamètre existant dans la partie basse du nuage (domaine du type ellipse 1 de la Figure 14, sur lequel les basses notes d'IBD ne peuvent pas être expliquées uniquement par les valeurs calculées du nutriment-cible de la relation).

Sur ces points donnant les valeurs d'indice anormalement basses, la plupart du temps, l'impact multiparamètre n'est pas douteux. Mais il est aussi envisageable que, sur certains points répartis de façon plus diffuse, mais plutôt en partie basse du nuage de points, la moyenne calculée à partir du résultat d'analyses chimiques donne une valeur sous-estimée par rapport à la valeur intégrée réelle des concentrations réellement subies *in situ* par les biofilms diatomiques (mauvaise prise en compte et non-intégration de pics de concentration par la stratégie d'échantillonnages ponctuels, ou plus généralement représentativité spatio-temporelle défectueuse de l'échantillonnage réalisé, conduisant à une mauvaise représentation de l'exposition moyenne subie).

Quoi qu'il en soit, ces premières analyses de relations soulignent à la fois **l'aspect pertinent**, mais aussi **les limites** des choix réalisés par le GIG Large Rivers pour la conduite de l'exercice d'intercalibration auquel la France a participé en 2016-2017. En effet, elles confirment bien qu'à cette échelle de taille de cours d'eau : 1) **le PO₄ reste bien l'un des paramètres d'altération les plus influents sur les flores diatomiques** (d'où la pertinence d'organiser l'exercice vis-à-vis de ce gradient) ; mais aussi 2) que la réponse des flores diatomiques à ce paramètre est altérée et se trouve fortement bruitée par **l'intervention multiple d'autres descripteurs d'altérations anthropiques** (d'où la qualité dégradée de la relation pression-impact obtenue en France, mais aussi à l'échelle européenne, sur ces grands cours d'eau).

Au final, toutes ces raisons conjuguées conduisent à l'expression de **relations pression-impact affaiblies** par rapport à des cours d'eau plus petits. Cependant, les résultats d'IBD₂₀₀₇ sur TGCE restent reliés de façon significative aux concentrations de PO₄. Par contre, dans le même contexte général d'altérations multiparamètres, les résultats de l'étude pression-impact concernant NH₄ sont révélateurs du fait que, le plus souvent, ce n'est visiblement pas le paramètre NH₄ qui occasionne l'effet biologique dominant à cette échelle de cours d'eau.

Comme attendu, il s'est confirmé que **l'assise temporelle prise en compte** influe sur la **qualité de la relation observée** entre les teneurs en nutriments et les notes d'IBD. En effet, malgré une réduction de l'assise-relevés et de l'assise de chimie correspondante, le raccourcissement du créneau temporel pris en compte améliore jusqu'à un certain point la qualité de la relation exprimée entre PO₄ et notes d'IBD (augmentation de la pente et du R²).

Ainsi, le passage d'un calcul moyen annuel à une intégration sur 4 mois (120 jours) et sur 3 mois (90 jours) améliore la qualité de relation, montrant qu'une intégration temporelle en gros équivalente à la durée de vie du biofilm en place donne l'image de l'exposition qui concorde le mieux avec l'impact biologique subi. Si un très léger avantage, perceptible seulement au niveau des valeurs non-arrondies, est encore décelé pour un raccourcissement de 120 à 90 jours, ces 2 temps d'intégration sont pratiquement équivalents au niveau pente et du R^2 de la relation, et la durée optimale d'intégration temporelle à adopter pour les diatomées se situe dans cette fourchette de durées.

Par contre, il n'est plus possible de considérer sérieusement le raccourcissement de la période prise en compte à 2 mois (60 jours). Outre une perte sèche de quelques relevés qui ne sont plus du tout documentés au niveau des conditions chimiques d'accompagnement, ce raccourcissement excessif provoque surtout une perte de représentativité et de robustesse de l'intégration chimique réalisée sur tout le jeu de données (2 relevés en moyenne, avec un effectif significatif de sites n'ayant plus qu'un relevé chimique disponible).

Ces résultats ont permis de confirmer qu'au moins pour les descripteurs de pression anthropique, il est préférable de représenter la qualité d'eau intégrée au site d'observation **de façon saisonnière** et en procédant à une intégration temporelle réalisant le **meilleur compromis** entre **chimie moyennée** et **épisode de croissance du biofilm prélevé**. De ce fait, la pratique adoptée pour la suite de cette étude, ainsi que pour la réalisation d'une autre étude nationale portant sur la détermination des caractéristiques autoécologiques des diatomées de France, a été d'adopter le pas de temps indiqué en haut de Figure 2, à savoir **une durée d'intégration de 90 jours** (soit approximativement 3 mois), déclinés par rapport au jour de prélèvement diatomique du relevé, avec l'application automatisée d'une **fenêtre temporelle [J-75 – J+15]**.

3.4.4. IBD₂₀₀₇ et nutriments (requêtes 2017, relations consolidées)

Après expertise complémentaire de données visant à adopter une pratique un peu plus ménagée pour la prise en compte de codes-remarques manquants ou erronés, au regard du contenu numérique indiqué comme résultat d'analyse du paramètre considéré, les relations actualisées entre nutriments et notes d'IBD, obtenues sur la base la plus complète possible disponible pour cette étude (requêtes complètes 2017), s'établissent comme suit.

La relation Pression-impact entre IBD₂₀₀₇ et orthophosphates est présentée en **Figure 17** ci-dessous :

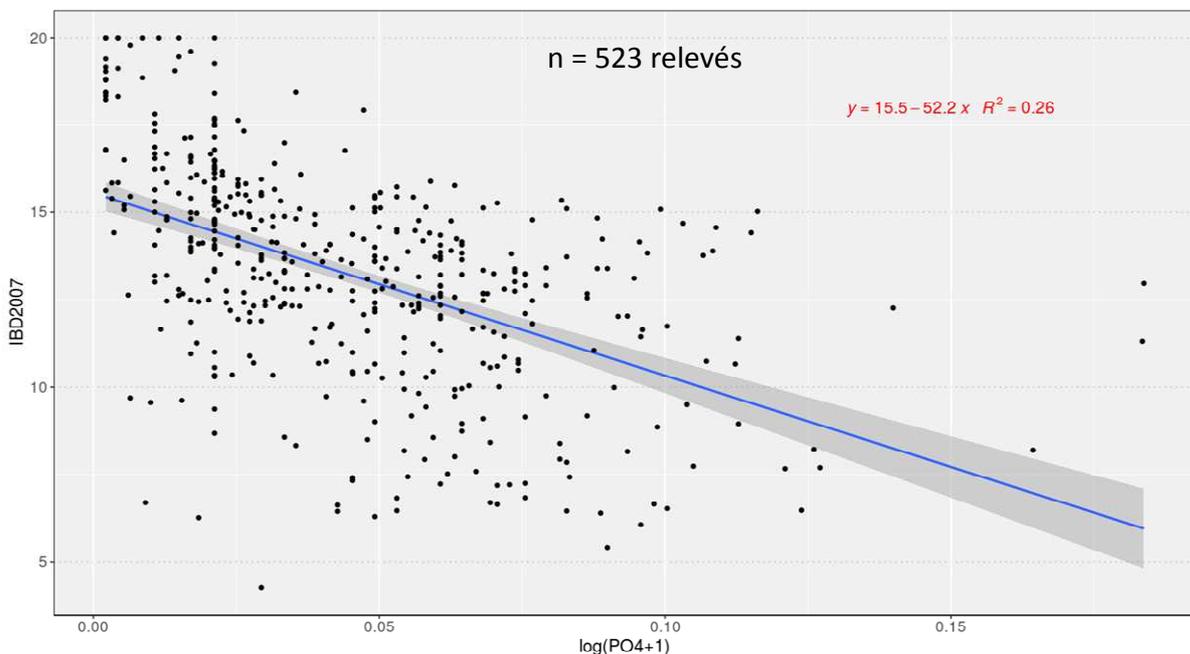


Figure 17 : Relation pression-impact entre IBD et PO₄ sur TGCE (intégration temporelle de 90 jours)

La relation linéaire obtenue après transformation log s'ajuste selon la formule suivante :

$$y = - 52 x + 15,5$$

Le R^2 de cette relation entre les notes d'IBD₂₀₀₇ et les concentrations de PO₄ s'établit à la valeur de 0,26. L'existence de cette relation sur le jeu de données national TGCE, bien qu'indéniable, est visiblement altérée et bruitée par un contexte multi-paramètres (points situés dans la partie la plus basse du nuage), ce qui lui fait visiblement intercepter l'axe des y à des valeurs trop basses d'IBD₂₀₀₇. Elle présente ainsi une qualité plus modeste que dans le cas du jeu de données national, tous cours d'eau confondus (cf. Figure 11, assortie d'un R^2 de 0,452).

La **Figure 18** présente la relation qui s'établit, sur la même base de 523 relevés TGCE, entre notes d'IBD₂₀₀₇ et concentrations en NH₄. Dans la configuration de ce jeu de données manifestement sous influence d'altérations composites, la relation obtenue présente une pente quasi-nulle, ce qui reviendrait à dire que ce paramètre n'exercerait pas d'influence structurante globale sur les notes d'IBD obtenues.

La réalité est plus complexe, car on voit bien que les valeurs du toit du nuage tendent vers une baisse avec l'augmentation des valeurs de NH₄ qui, dans cette position, sont plus probablement illustratives d'un contexte où la pression de NH₄ est dominante, voire monofactorielle. Mais par rapport à des cours d'eau plus petit, le gradient de NH₄ est ici nettement réduit (valeurs les plus fortes n'atteignant pas 0,3 alors que, dans la même échelle transformée log (cf. Figure 12), le gradient représenté montait jusqu'à des valeurs de plus de 2. D'autre part, dans ce jeu de données TGCE, une donnée extrême (concentration transformée log proche de 0,3) semble entachée d'un important défaut de représentativité temporelle. Il a dû être échantillonné un pic très transitoire de concentration de NH₄ qui conduit à un résultat de concentration moyenne élevée, probablement non-représentative de la concentration réelle en NH₄ intégrée *in situ* (d'où une réponse indicelle nettement moins affectée qu'attendu dans un tel cas). Ce cas particulier très visible, entre autres plus inclus dans le nuage de points, donc moins visibles, est illustratif du défaut de représentativité temporelle d'échantillons ponctuels pour représenter la chimie de l'eau. Par ailleurs, du fait de difficultés plus importante d'accessibilité dans le lit, de tels grands systèmes sont souvent plus difficiles à échantillonner de façon représentative, car ils sont souvent sujets à des phénomènes d'influences localisées liées à des apports de bordure et à une plus grande difficulté d'homogénéisation rapide de la qualité de la colonne d'eau (cf. importance de la section mouillée).

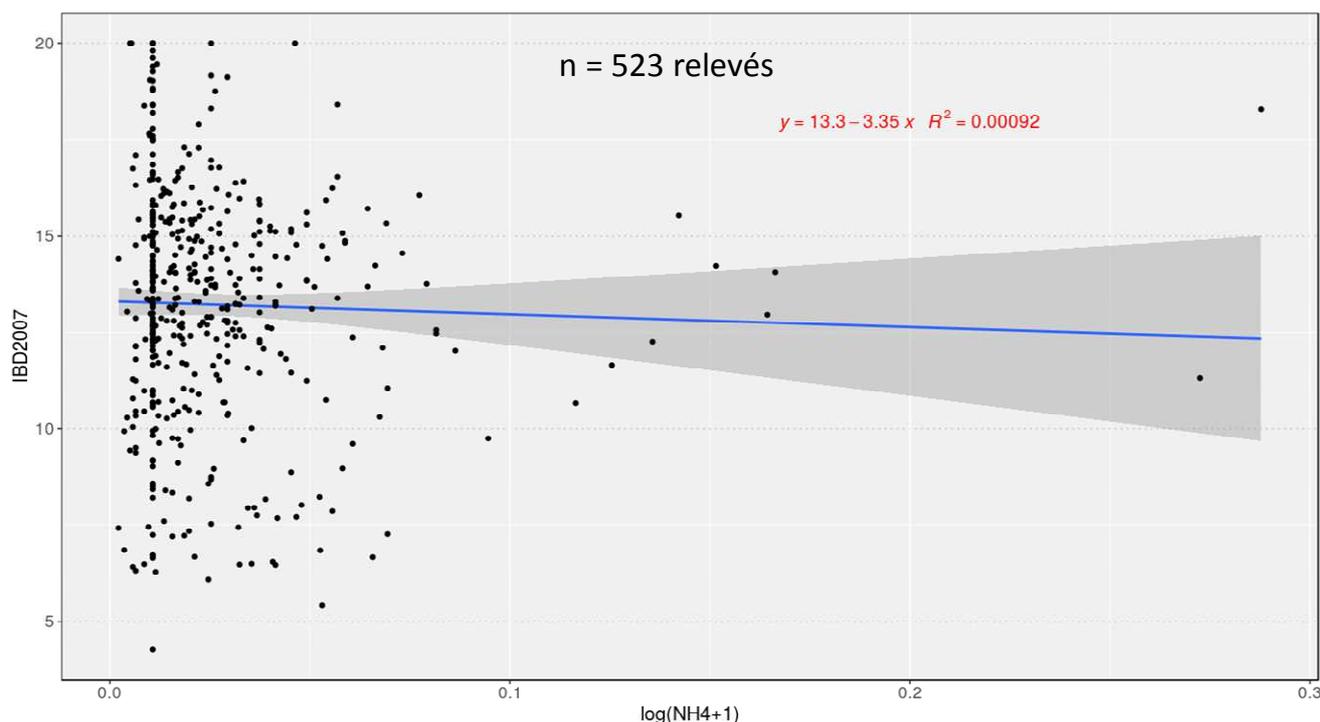


Figure 18 : Relation pression-impact entre IBD et NH₄ sur TGCE (intégration temporelle de 90 jours)

Enfin, beaucoup de valeurs indicelles basses à très basses (cf. positionnement analogue à celui de l'ellipse 1 dans la Figure 14) sont obtenues dans un contexte de relativement faibles valeurs de NH₄, et même de valeurs de NH₄ non quantifiées (post-calcul fait à partir des seuils analytiques de quantification non-atteints).

Dans de telles situations, il est patent que l'on se situe dans un contexte de multi-altération et que l'impact biologique dominant se répercutant sur la note d'IBD est certainement lié à d'autres types d'altérations anthropiques que le NH₄.

Quoi qu'il en soit, cette relation linéaire défectueuse illustre bien le fait que, en TGCE et pour différentes raisons évoquées précédemment (plus faible amplitude du gradient présent, intervention de la dilution et de processus auto-épuratoires, difficultés accrues de représentativité d'échantillonnages ponctuels sur d'aussi grands systèmes etc...), la pression liée à l'ammonium n'est pas dominante et a souvent tendance à être masquée par l'intervention d'autres paramètres plus influents.

Pour les autres descripteurs physico-chimiques d'altération anthropique (Sat O2, DBO5, NO3 etc...), la diminution de qualité de régression avec les notes d'IBD2007 sur TGCE, par rapport à ce qui se passe sur un jeu de données tous cours d'eau confondus, est aussi la règle et le résultat oscille selon les descripteurs entre ce qui se passe pour PO4 (relation atténuée mais restant malgré tout démontrée) et ce qui se passe pour NH4 (facteur d'influence très minoré dans un tel contexte, aboutissant à une relation pression-impact de mauvaise qualité).

3.5. Relation pression-impact sur la base d'un gradient composite TGCE

Afin de glaner des éléments plus précis sur la hiérarchisation des raisons occasionnant une dégradation des relations pression-impact en très grands cours d'eau, nous avons étudié plus avant l'influence de cette grande taille de système afin de voir si elle était elle-même génératrice d'un décalage par rapport à la méthodologie générale utilisée pour la mise au point de l'IBD₂₀₀₇.

Le calage initial de l'IBD₂₀₀₇ (cf. REF. 7) a été effectué sur la base d'une ACP mobilisant, sur un jeu de données national de 2 802 relevés diatomiques disponibles à l'époque (toutes tailles et tous types de cours d'eau confondus), 7 descripteurs abiotiques significativement impactés par les altérations anthropiques (voir Figure 19, partie gauche).

Un **gradient composite d'anthropisation**, construit à partir de la contribution de ces 7 variables sur 2 axes de l'analyse représentatifs des gradients de pollution (Axe 1 plutôt représentatif de pollutions composites d'origine domestique et/ou de STEP associant nutriments et altération organique ; Axe 3 représentatif de pollutions agricoles diffuses par les nitrates) et des valeurs de paramètres chimiques associés à chaque relevé diatomique, a permis de générer 7 classes de qualité d'eau (partie droite de la figure). Chaque point correspond au positionnement dans le gradient composite d'un relevé élémentaire (doublet données physico-chimiques-relevé diatomique correspondant).

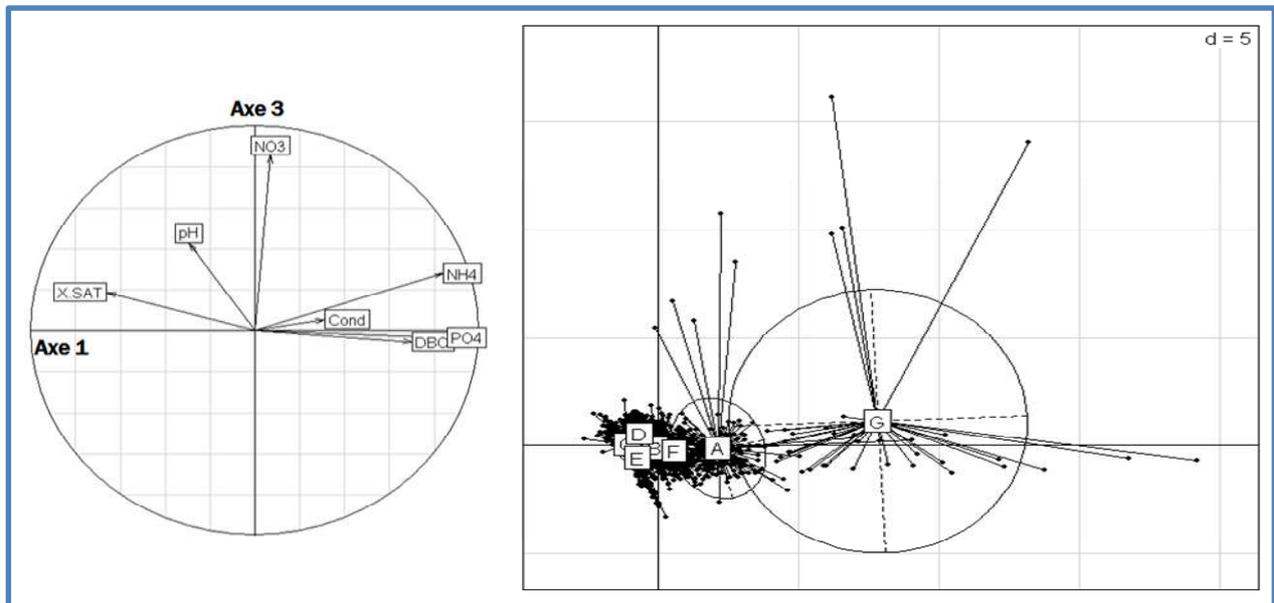


Figure 19 : ACP sur 7 descripteurs d'anthropisation, utilisée pour la genèse de l'IBD₂₀₀₇ (partie gauche) ; détermination des 7 classes de qualité correspondantes (partie droite).

Les profils indiciels des taxons contributifs de l'IBD ont ensuite été construits sur la base de la probabilité de présence d'un taxon donné dans ces 7 classes de qualité, en repartant des valeurs physico-chimiques associées à chaque relevé contenant ce taxon et en les convertissant en valeur obtenue dans l'échelle de gradient composite.

Le constat ayant été fait que les régressions pression-impact obtenues sur TGCE vis-à-vis de paramètres physico-chimiques élémentaires sont affaiblies par rapport à celles obtenues sur des cours d'eau plus petits, une des raisons explicatives pourrait tenir à l'hypothèse que l'équilibre relatif des paramètres abiotiques ayant participé au calage initial de l'IBD₂₀₀₇, tous types de cours d'eau confondus, serait nettement modifié en cas de TGCE (pollutions composites intervenant dans un équilibre différent à cette échelle de taille de cours d'eau).

Pour confirmer ou infirmer cette hypothèse, il a été tenté de reconstruire un gradient composite d'anthropisation avec la même méthodologie que celle utilisée pour créer l'IBD₂₀₀₇, mais en travaillant à partir du jeu de données TGCE disponible (543 relevés de sites d'observation > 8 000 km²), puis d'étudier la qualité de la relation pression-impact obtenue sur cette base.

De façon analogue à la démarche antérieure d'élaboration de l'IBD₂₀₀₇, il a été réalisé une ACP construite à partir des 7 descripteurs abiotiques de pression anthropique associés aux relevés TGCE disponibles (partie gauche de la **Figure 20**), pondérés par leur contribution issue de l'IPS (colonne de gauche du **Tableau 3**).

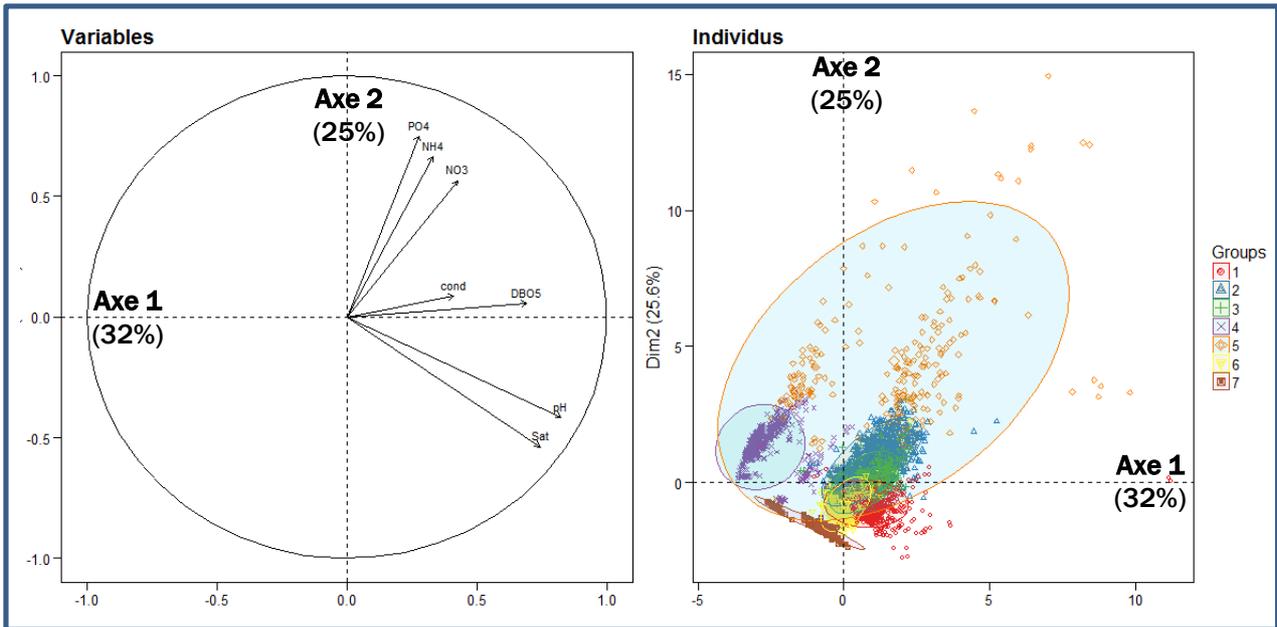


Figure 20 : ACP réalisée à partir des 7 descripteurs d'anthropisation de l'IBD₂₀₀₇, sur jeu de relevés nationaux strictement TGCE (partie gauche) ; détermination des 7 classes de qualité correspondantes (partie droite).

Par rapport à l'approche initiale tous cours d'eau confondus, on peut remarquer (cf. **Figure 20**, partie gauche) que la structuration de l'ACP sur TGCE réalisée sur données moyennées sur 90 jours (J-75 – J+15) répond à un patron sensiblement différent, s'organisant de façon très dominante selon les 2 premiers axes de l'ACP, qui portent une contribution cumulée conséquente (57 %).

Tableau 3 : Colonne centrale : pondération appliquée aux paramètres IBD (selon niveau de contribution calculé dans l'IPS); à droite : poids relatif issu de l'ACP pour le calcul du gradient composite d'altération

Pondération	IPS	Poids
pH	0,003	0,01
Conductivité	0,046	0,08
% saturation	0,099	0,17
DBO5	0,077	0,13
NH4	0,144	0,24
NO3	0,05	0,08
PO4	0,172	0,29
Total	0,591	1

Dans le cadre de cette ACP TGCE, l'Axe 1 est plus porteur du gradient de pollution organique et de conductivité. L'Axe 2, pour sa part, est porteur d'une opposition de contribution entre saturation de l'eau en O₂ (vers le bas, signe des meilleures qualités rencontrées dans ce type de grands hydrosystèmes) et nutriments dissous vers le haut (NH₄, PO₃, NO₃).

L'influence représentée sur l'axe 2 serait plutôt représentative de pollutions diffuses d'aval de bassins versants et de rejets de STEPs ayant déjà subi une certaine minéralisation de la matière organique (en liaison avec l'augmentation des temps de résidence et la capacité auto-épuratrice propres à de tels hydrosystèmes ?).

En tout cas, cette différence assez importante de physiologie générale de l'ACP montre bien que les paramètres abiotiques d'altération n'interviennent ni avec la même intensité (fait vérifié par ailleurs en fonction de l'ampleur réduite des gradients, qui se répercute sur les scores de gradient composite), ni dans le même équilibre que pour le jeu de données constitué d'hydrosystèmes plus petits.

Les 7 classes de qualité telles que représentées en partie droite de la Figure 20 ont été obtenues suite à une classification basée sur la Méthode de Ward. On peut là aussi remarquer qu'elles obéissent à un patron d'organisation apparemment différent par rapport au jeu de données basé sur des cours d'eau plus petits (partie droite de la Figure 19). Le gradient composite semble se distribuer selon 2 alignements différents, probablement en fonction de 2 grandes typologies dominantes de pollutions rencontrées. Les classes déterminées en fonction des scores de gradient composite interceptent plus ou moins en parallèle ces 2 alignements différents (le gradient s'établissant à peu près en diagonale des 2 axes), avec les sites de la meilleure qualité (ici, Cluster 7) situés en bas et vers la gauche du nuage et les sites les plus altérés (ici Cluster 5) s'étirant vers le haut et la droite.

Dans la suite logique de cette démarche, nous avons ensuite étudié la relation pression-impact établie à partir des scores de gradient composite et des notes d'IBD par site, présentée en **Figure 21** sur le créneau d'intégration temporelle déjà établi comme le plus pertinent pour ce maillon des biofilms diatomiques benthiques (soit 90 jours).

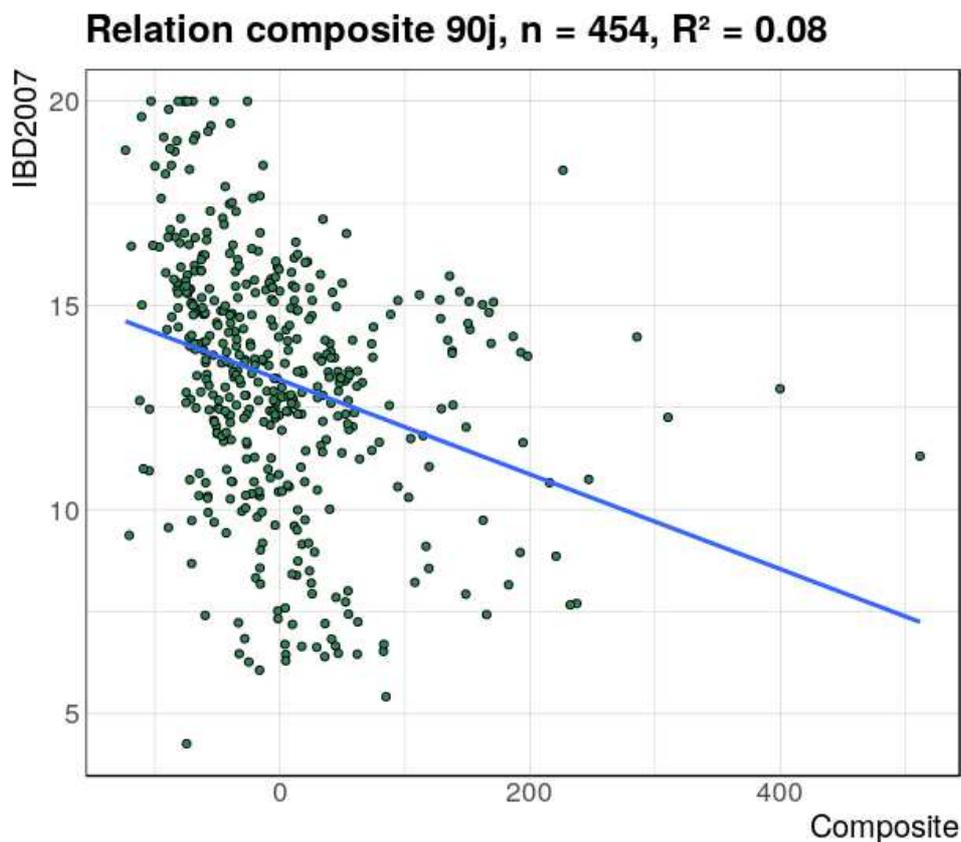


Figure 21 : Relation pression-impact établie sur TGCE entre scores de gradient composite et notes d'IBD₂₀₀₇

Ce graphique est à nouveau révélateur d'une relation très probable (cf. physiologie du toit du nuage et pente décroissante des notes d'IBD en fonction de l'augmentation des valeurs de score de gradient composite obtenues), mais aussi très floutée. Comme pour les études de nutriments séparés et alors que la prise en compte des 7 paramètres d'altération aurait dû améliorer la qualité de la relation, il existe un niveau élevé de résidus par rapport à l'ajustement linéaire modélisé et le R² s'établit au niveau très médiocre de 0,08.

Le fait de prendre en compte, via le gradient composite, l'influence cumulée des 7 descripteurs abiotiques d'altération anthropique n'a pas résolu, ce qu'il aurait dû contribuer à faire, l'aspect de la forme en fanion du nuage de points, l'ajustement discutable de la régression linéaire obtenue et le constat que beaucoup de réponses biologiques visiblement impactées ne sont pas expliquées, sur les cas particuliers, par l'intervention de paramètres d'altération captés par l'analyse chimique, alors que le gradient composite a été construit à partir de la contribution de ces paramètres à l'échelle de tout le jeu de données. Par contre, la physionomie du toit du nuage et la pente que présente la relation semblent à mettre au crédit de l'existence de cette relation.

L'hypothèse la plus vraisemblable à ce stade pour expliquer l'origine du problème rencontré, au moins dans le cas de cette figure, est très probablement qu'une stratégie de description de l'ambiance chimique intégrée basée sur seulement 2 ou 3 analyses ponctuelles présente un **défaut de représentativité spatio-temporelle encore plus important sur TGCE que sur des hydrosystèmes plus petits**. Ainsi, de nombreuses réponses biologiques situées dans le bas du nuage et plus particulièrement dans le secteur de l'ellipse 1 (cf. Figure 14) sont révélatrices d'impact biologique indéniable capté par l'IBD, alors que la contamination physico-chimique du milieu aquatique en nutriments, **pourtant restituée de façon intégrative** sur la base d'un **gradient composite** construit à partir des valeurs des 7 paramètres abiotiques ayant servi à caler l'IBD, ne semble pas pouvoir justifier l'obtention de notes si basses. Alors que, pour les relations avec le PO₄ et le NH₄ présentées par exemple en Figures 17 et 18, il était possible de faire l'hypothèse que les basses valeurs indicielles obtenues dans une zone de faibles valeurs du paramètre considéré soient dues à l'action conjointe d'autres descripteurs d'altération anthropique que celui faisant l'objet de la représentation graphique, cette hypothèse de travail ne peut plus être légitimement évoquée dans ce cas (le gradient abiotique représenté est composite) et il devient assez évident que, dans le cas des TGCE et à partir d'échantillonnages ponctuels, l'ambiance chimique trophique est le plus souvent mal décrite et probablement sous-estimée.

Donc, sauf à affirmer que l'IBD répondrait de façon erratique et déconnectée de l'ambiance intégrée en nutriments de la colonne d'eau, alors que son efficacité pour révéler l'impact de telles pollutions a été largement démontrée et publiée par ailleurs, il se confirme ici que le **caractère bio-intégrateur du maillon diatomique phytobenthique est largement plus robuste et représentatif** du degré d'altération physico-chimique (volets des pollutions organiques et de l'enrichissement en nutriments) que la **stratégie de prélèvements ponctuels** censés représenter fidèlement la qualité de l'eau, et que **ce défaut de représentativité s'aggrave sensiblement en TGCE** par rapport à des hydrosystèmes de taille inférieure.

Ces résultats, qui rejoignent le constat collectif de perte de qualité de relation des indices nationaux et de l'ICM vis-à-vis des nutriments lorsqu'on s'intéresse au cas des Large Rivers, viennent indirectement conforter la pertinence de la stratégie de surveillance de l'état écologique mise en place par la DCE, qui a recommandé de s'appuyer sur différents maillons-clés bio-intégrateurs afin d'évaluer les atteintes à l'état écologique de façon plus intégrative et plus robuste qu'avec les seules analyses chimiques réalisées sur ces milieux.

Devant ces résultats existants mais qui, en fonction des raisons pré-évoquées, conduisaient à l'illustration de relations pression-impact d'un niveau plutôt atténué sur cette typologie de cours d'eau, nous avons tenté d'obtenir des résultats complémentaires permettant de répondre de façon plus complète à la sollicitation du GIG, qui nous avait demandé d'étayer et de défendre la validité et la pertinence, dans ce cadre d'application TGCE, de notre système national d'évaluation basé sur l'IBD₂₀₀₇. Un autre volet d'établissement de relations pression-impact a donc été ensuite abordé. Il a consisté à étudier et à formaliser les relations qui pouvaient s'établir entre les traits auto-écologiques des assemblages diatomiques et l'évaluation d'altération réalisée par notre système indiciel national (voir chapitre suivant).

4. Liens entre IBD₂₀₀₇ et attributs autoécologiques des communautés

Cette mise en relation a été effectuée en utilisant des bibliothèques de traits auto-écologiques des espèces de diatomées d'eau douce mises au point par van Dam à partir de jeux de données diatomiques collectés aux Pays-Bas (REF. 8 : van Dam, H., 1994) dans le cadre de différentes études scientifico-techniques et de suivis de réseaux de surveillance. En pratique, les bibliothèques de traits qui ont été utilisées pour la réalisation de ce travail sont celles contenues et mises à jour par Michel COSTE (Irstea) dans le logiciel OMNIDIA sur la base de cette publication.

De telles approches permettent de tester et de vérifier l'existence de relations pression-impact à partir de l'exploitation d'informations intégratives d'origine biologique captées au niveau des assemblages diatomiques (composition de la communauté résumée par sa composition en classes auto-écologiques plutôt que par sa composition spécifique détaillée).

4.1. Liens autoécologiques IBD₂₀₀₇ / trophie (PO₄), tous types de cours d'eau confondus

En 2007, alors que le nouvel indice IBD₂₀₀₇ venait d'être créé, nous avons dû légitimer son entrée en service par rapport à la version en vigueur à l'époque (IBD₂₀₀₀) en démontrant la plus-value apportée par ce nouvel outil, qui devait être porteur d'un niveau de performance amélioré et/ou permettre une évaluation régionalisée plus judicieuse.

Ces démonstrations et justifications ont dû être réalisées successivement à 2 niveaux :

- au niveau national (Direction de l'Eau et ONEMA tout juste créé), préalablement à la publication d'une Circulaire Nationale Evaluation faisant passer cet indice en utilisation de routine, à la place de l'IBD₂₀₀₀, dans le cadre des réseaux nationaux d'évaluation ;
- au niveau Européen, dans le cadre des exercices d'intercalibration menés par 3 groupes géographiques d'intercalibration ou GIGs (Alpin, Central-Baltique, Méditerranéen, selon la zone géographique des sites nationaux concernés), en vue de démontrer la DCE-compatibilité de l'évaluation nationale basée sur cette nouvelle version d'indice.

Outre la mise en évidence de l'amélioration de relations pression-impact classiques, basées sur des données de chimie de l'eau (cf. notamment les Figures 11 et 12 du présent rapport), une autre voie utilisée pour défendre notre nouveau dispositif national d'évaluation a visé à vérifier la cohérence entre la classification d'état écologique obtenue de façon régionalisée à partir des notes d'IBD₂₀₀₇ et la répartition des caractéristiques auto-écologiques des communautés multispécifiques basée sur la classification proposée par van Dam. En 1994 en effet, cet auteur avait classé les espèces diatomiques connues à l'époque aux Pays-Bas dans différentes classes correspondant à leur degré de sensibilité-tolérance à un gradient-environnemental-cible donné (ex ; pH, saturation de l'eau en O₂, minéralisation de l'eau, trophie, N-trophie...). Ainsi, les taxons étaient répartis dans des classes autoécologiques s'échelonnant depuis les sites de référence jusqu'aux plus fortes valeurs du gradient environnemental considéré (soit, pour des paramètres représentatifs d'altérations anthropiques, depuis le très bon état jusqu'à la plus mauvaise classe d'état). Le nombre de classes était variable selon le paramètre environnemental considéré et selon l'amplitude du gradient représenté aux Pays-Bas pour ce paramètre.

Dans le cadre de la validation de performance de l'IBD₂₀₀₇ indiquée plus haut, les tests destinés à décrire les communautés à partir des traits autoécologiques ont été réalisés sur le jeu de données national tout-venant disponible à l'époque, qui comportait 845 échantillons couplant biologie et chimie complète, prélevés dans tous les types de cours d'eau (des très petits aux très grands).

La **Figure 22** présente le pourcentage d'espèces sensibles (somme des espèces oligotrophes + mésotrophes) dans des conditions de référence et au long du gradient d'état écologique.

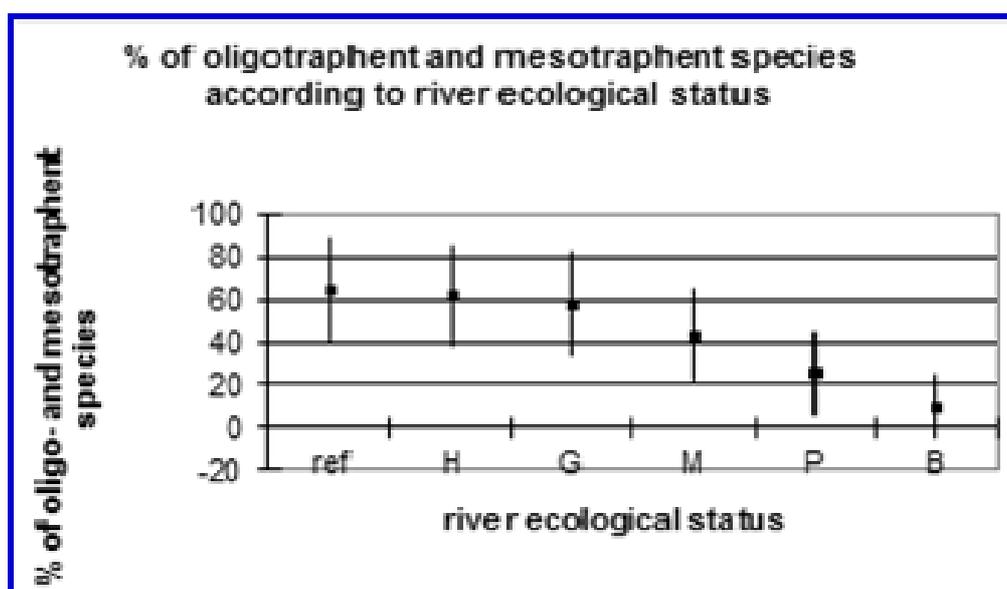


Figure 22: Caractéristiques trophiques des communautés diatomiques de France sur sites de référence et au long du gradient de classes de qualité des eaux (depuis le très bon état jusqu'au mauvais état)

Sur la base de classes de qualité régionalisées établies à partir d'EQRs-IBD₂₀₀₇, ce graphique montrait :

- l'absence de différence significative en pourcentage d'espèces sensibles entre l'état de référence et le très bon état (high status ou H);
- Une très légère diminution, mais cependant significative, des espèces sensibles entre classe de très bon état (H) et classe de bon état (G). Néanmoins, l'abondance relative moyenne des espèces sensibles reste dominante.
- Une baisse du pourcentage d'espèces sensibles entre état bon (G) et moyen (M), illustrant un net changement de structure des communautés en réponse à la dégradation anthropique de l'environnement aquatique.

Sur la base de cette analyse, la limite entre le bon état et l'état moyen telle qu'elle est appliquée en France apparaît bien calée, en accord avec la définition normative du bon état dans la DCE.

4.2. Liens IBD₂₀₀₇ / trophie sur une sélection de Très Grands Cours d'Eau

Dans le cadre de l'exercice d'intercalibration auquel la France a participé en 2015-2016, le GIG « Large Rivers » a sollicité l'illustration de relations pression-impact si possible spécifiquement établies sur les Très Grands Cours d'Eau, et pouvant venir à l'appui d'une démonstration de la performance d'évaluation des systèmes nationaux d'évaluation proposés par les Pays-Membres participants (dont la France) sur de tels hydrosystèmes.

En procédant d'une façon comparable à la méthodologie analytique mise en œuvre en 2007 sur un jeu Français tous cours d'eau confondus (cf. Figure 22), nous avons donc entrepris, à partir de la requête la plus complète qu'il était possible de mobiliser en début de 2^{ème} semestre 2015, un travail d'analyse de données visant à vérifier la cohérence entre la classification biologique donnée par l'IBD₂₀₀₇ et la composition des communautés en termes de caractéristiques autoécologiques. Du fait que cette analyse ne mobilisait que des informations biologiques issues de la composition des relevés diatomiques, il a été possible de s'appuyer sur un sous-jeu de données important, composé de 645 échantillons de diatomées uniquement collectées sur des Très Grands Cours d'Eau (seuil de sélection basé sur des bassins versants intégrés > 8 000 km²).

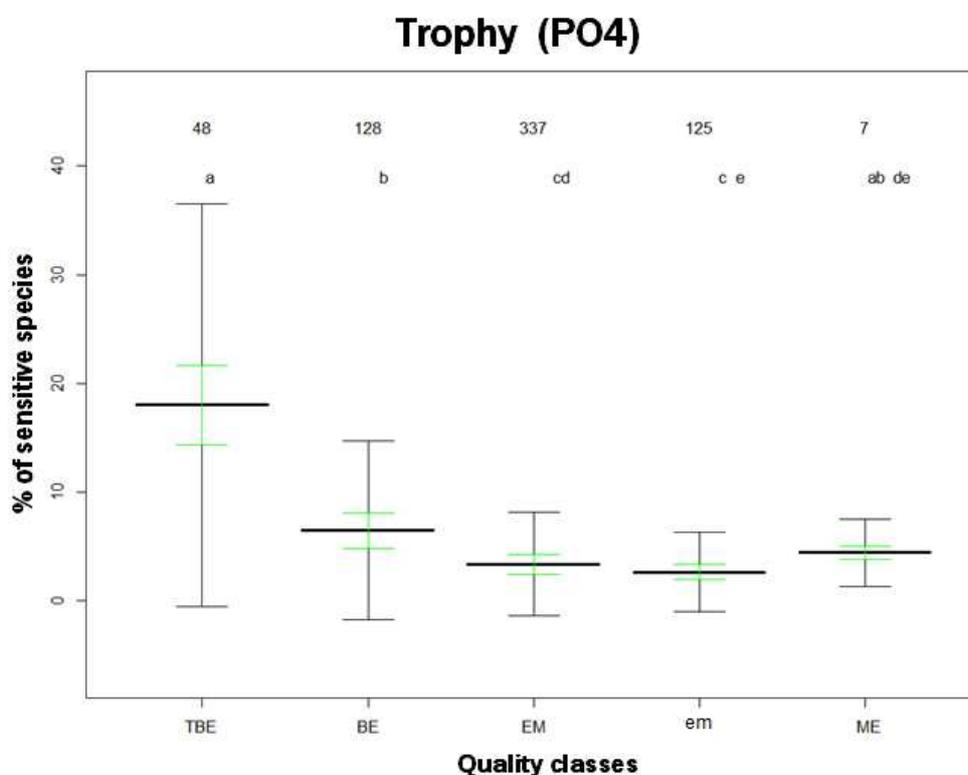


Figure 23: Pourcentage d'espèces sensibles au PO₄-par classe de qualité en TGCE (> 8 000 km²)

En France, du fait de l'anthropisation omniprésente, il n'existe plus aucune situation de référence valide sur TGCE, et il n'est donc pas possible de présenter et de calculer une information représentative de ce statut. La **Figure 23** ne présente donc que les 5 classes classiques de qualité des eaux utilisées dans le cadre d'application de la DCE. Outre cette différence de présentation, la comparaison entre Figure 23 et Figure 22 montre une répartition d'attributs autécologiques obtenus par classe de qualité sensiblement différente par rapport aux cours d'eau tous types confondus.

Même si la tendance générale obéit en gros au même patron de structuration (plus grande abondance des taxons sensibles dans les classes de meilleures qualités, puis diminution pour les classes subissant des niveaux croissants d'altération), il est à noter que, pour les classes de TBE et de BE, les niveaux de taxons sensibles à la trophie sont sensiblement plus bas que pour l'ensemble de données national antérieur intégrant tous les types de cours d'eau.

Une première raison tient sans doute au fait que, pour les TGCE, plus aucun site de référence réel ne subsiste, et sans doute aussi au fait qu'une proportion inférieure de sites de très bon état existe, principalement distribuée dans la moitié basse de la classe (donc avec une répercussion sur les attributs autoécologiques conduisant à une réduction du pourcentage de taxons sensibles au PO₄).

Une autre source possible de cette réduction des taxons sensibles sur d'aussi grands hydrosystèmes peut trouver son fondement dans le concept de continuum fluvial proposé par Vannote et al, 1980 (**REF. 9**). En effet, cet auteur émet l'hypothèse forte que les caractéristiques structurelles et fonctionnelles des communautés dans les rivières sont adaptées pour se conformer à la localisation la plus probable ou à l'état moyen du système physique au site.

Dans les hydrosystèmes naturels, la traduction de ce concept signifie que les conditions physiques et l'état trophique se situent à un niveau différent à l'amont et à l'aval, avec une augmentation normale du niveau naturel de trophie vers l'aval. Dans notre jeu de données national seulement composé de TGCE, plus aucun d'entre eux n'étant indemne d'influence anthropique, cette tendance naturelle est amplifiée par l'influence de diverses sortes de pressions anthropiques (par exemple la pollution agricole diffuse, les effets cumulés des grandes villes et des stations d'épuration le long des rivières ...). C'est probablement l'une des raisons pour lesquelles nous pouvons observer, sur ce jeu de données TGCE, une telle différence de niveau des attributs auto-écologiques trophiques des communautés par classe de qualité, par rapport au jeu de données national de 2007 qui englobait toutes les tailles de cours d'eau.

Nous avons aussi détecté un artefact méthodologique qui peut contribuer assez fortement à cet écart constaté. Dans sa classification de trophie, van Dam a « défini » une Classe 7 de trophie très peu précise et très large (couvrant du domaine oligotrophe au domaine hypereutrophe) incluant des taxons qui n'avaient pas pu faire l'objet d'une classification précise en termes de trophie dans les conditions des Pays-Bas. La raison (non précisée pour chaque taxon) pouvait en tenir, soit à une plasticité écologique vraiment large de l'espèce vis-à-vis du paramètre considéré (*i.e.* taxon très euryèce), ce qui peut effectivement exister ; soit aussi, ce qui est plus gênant ici et peut venir fausser les résultats présentés, à une présence tellement marginale du taxon dans le jeu de données d'origine (en termes d'occurrence et/ou en termes d'abondance relative dans les relevés) qu'il n'a pas été possible à Herman van Dam de lui attribuer une valeur auto-écologique précise.

Même si, pour l'élaboration de la Figure 22 comme de la Figure 23, il a été procédé de la même façon en regroupant ces taxons éventuellement mal connus de la Classe 7 avec les espèces tolérantes, l'assise-taxons de la Classe 7 a pu influencer, de façon différenciée selon ces 2 cas, sur la physionomie des caractéristiques autoécologiques générales des communautés, d'autant que l'assise de taille des cours d'eau est franchement différente entre ces 2 jeux de données (toutes tailles de cours d'eau confondues en 2007 avec une forte dominance de petits à moyens cours d'eau, sélection uniquement basée sur de Très Grands Cours d'Eau en 2015, suivant en cela la requête du GIG).

Il n'a malheureusement pas été possible de ré-accéder facilement au jeu de données d'origine, ni aux tables exactes d'attributs de taxons d'Omnidia utilisés en 2006-2007, dans un contexte où la taxonomie a assez largement évolué depuis la réalisation de ces travaux, avec des regroupements et des séparation *pro parte* de taxons, et où les tables de traits ont été assez largement retouchées en fonction de ces modifications taxonomiques...).

De ce fait, il ne nous a pas été possible de réaccéder, sur exactement la même base d'attributs-taxons, à la ventilation des taxons de Classe 7 entre classes de qualité écologique telle qu'elle existait en 2007. Par contre, nous avons pu étudier la répartition des espèces indéterminées (Classe 7) selon les classes de qualité sur le jeu de données TGCE 2015. Cette ventilation est présentée en **Figure 24** (page suivante).

Trophy (PO4)

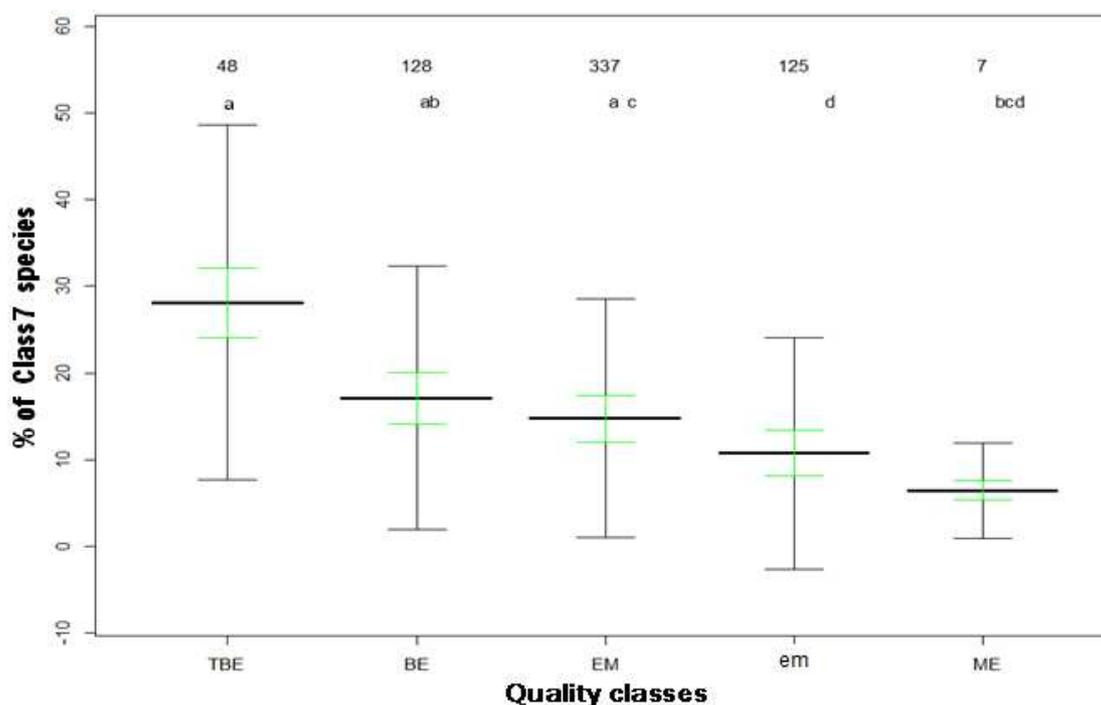


Figure 24 : Distribution des espèces de la Classe 7 (mal caractérisées vis-à-vis de la trophie) au long du gradient des 5 classes de qualité des Très Grands Cours d'Eau

On peut aisément constater qu'une quantité très substantielle de taxons indéterminés ou mal déterminés vis-à-vis de la trophie existe dans notre jeu de données TGCE. De plus, il apparaît nettement que leur distribution n'est pas équilibrée entre les classes, mais est très sensiblement plus forte dans la classe TBE (> 28% en moyenne) et BE (> 17% moyenne), que dans les classes d'état plus moyennes ou dégradées.

Malheureusement, ce gradient déséquilibré d'espèces mal déterminées par van Dam sur le plan de leur trophie fonctionne en sens inverse du gradient d'espèces sensibles, qu'il vient mécaniquement diminuer (puisque ces taxons à écologie mal déterminée ne peuvent pas être valablement considérés comme taxons sensibles, donc viennent par défaut se cumuler à l'ensemble des taxons tolérants et/ou d'écologie mal cernée). Ce gradient inverse vient donc impacter fortement la physionomie des résultats présentés dans la Figure 23, puisque ces taxons à l'écologie mal définie pour les Pays-Bas, rattachés ici à l'ensemble des taxons tolérants, ont pour effet de diminuer artificiellement et fortement le pourcentage des taxons sensibles à la trophie dans les classes de TBE et de BE.

La physionomie générale de la Figure 24 suggère que ce problème doit exercer une contribution importante à la baisse de niveau constatée des taxons sensibles des classes TBE et BE entre la Figure 22 (tous cours d'eau, 2007) et la physionomie de distribution des taxons sensibles de la Figure 23 (nouvelle analyse 2016 réalisée uniquement sur TGCE).

La raison écologique sous-jacente tient probablement au fait qu'aux Pays-Bas, en fonction des différences de types et d'altitude de cours d'eau, van Dam n'a jamais trouvé, ou n'a trouvé que de façon trop faible et éparse la plupart des espèces contribuant aux communautés typiques de très grands cours d'eau de France en très bon ou en bon état (les grandes rivières de la meilleure qualité étant la Dordogne et la Durance, qui arrivent assez directement de zones de haute altitude, respectivement du Massif Central et des Alpes). A l'inverse, il apparaît qu'il a pu qualifier l'autécologie de la plupart des taxons tolérants que nous trouvons aussi en France (beaucoup moins d'indéterminés de Classe 7 dans ces classes d'état plus dégradées), illustrant indirectement le fait que cette catégorie d'espèces est plus facile à trouver et/ou mieux représentée dans les hydrosystèmes des Pays-Bas.

Pour résoudre ce problème, la seule solution efficace serait de réitérer le même type d'analyse autoécologique que celle effectuée par van Dam en 1994 aux Pays-Bas, sur un jeu de données diatomiques le plus complet possible (toutes tailles et tous types naturels de cours d'eau confondus), collecté par les réseaux français de surveillance dans les contextes naturels diversifiés de notre propre territoire national.

Si notre équipe était consciente depuis assez longtemps de l'intérêt de réaliser un tel travail dès que possible, il a par contre été inenvisageable de le réaliser dans le temps contraint du démarrage de l'exercice d'intercalibration et de présenter rapidement les résultats modifiés des analyses qui précèdent ; une première raison tenant au fait que, suite à une requête large, le jeu de données chimique national nécessite tout d'abord un post-traitement initial très conséquent, qui nécessite du temps, avant d'être utilisable à cet effet.

Nota : Un travail de redéfinition des attributs autoécologiques des espèces de diatomées françaises, basé sur une assise-données la plus large possible, a depuis été entrepris par notre équipe et est actuellement en cours de réalisation. Il permettra de couvrir et de bien décrire des flores venant de contextes naturels largement plus diversifiés qu'aux Pays-Bas (qui est un pays plat et à substrats dominants acides, alors qu'en France, les substrats calcaires dominent...), et d'en déduire les caractéristiques autoécologiques d'un nombre sensiblement accru d'espèces, en meilleure correspondance avec la diversité des contextes biogéographiques de France et le cadre national d'application qui pourra en être fait (ex : métriques de diagnostic).

Du fait de ce problème, nous avons aussi essayé de voir quelle physionomie pourraient présenter les résultats en adoptant un principe inverse de construction de graphique, c'est-à-dire en axant la présentation de la **Figure 25** ci-dessous sur l'augmentation des espèces tolérantes à la trophie (PO4), incluant les espèces indéterminées de la classe trophique 7, en fonction de la dégradation des classes de qualité d'eau.

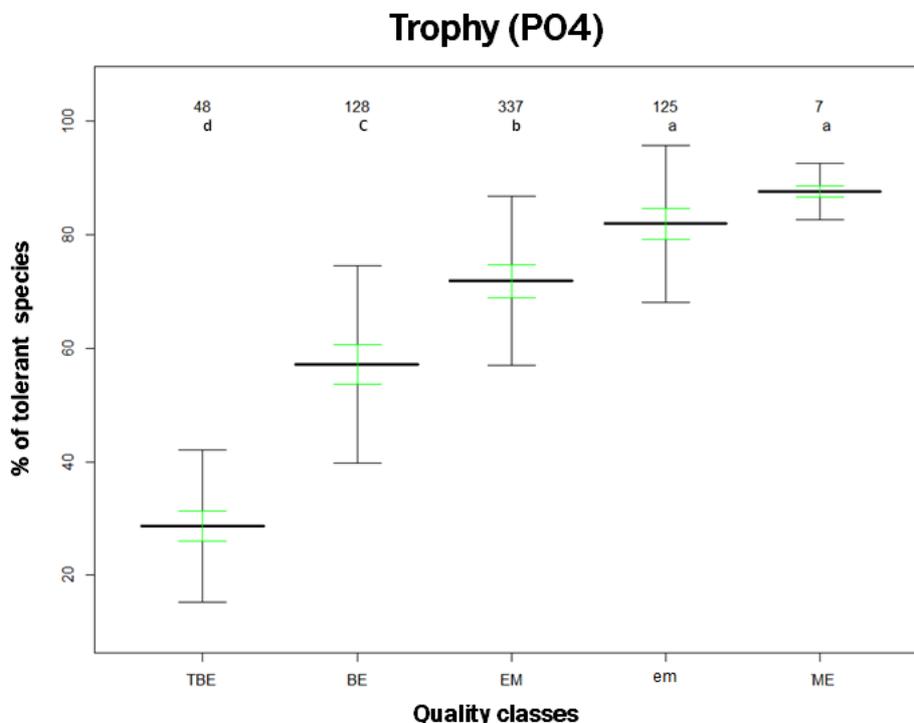


Figure 25: Pourcentage d'espèces tolérantes au PO4 par classe de qualité en TGCE (> 8 000 km2)

Même si un pourcentage important d'espèces au statut trophique souvent mal défini (cf. Figure 24) a été cumulé par construction avec les espèces tolérantes à la trophie, pratique qui conduit très probablement à surestimer le pourcentage d'espèces vraiment tolérantes dans les classes de TBE et de BE (puisque beaucoup d'indéterminées actuelles dans les classes de TBE et de BE devraient plutôt être des espèces oligotrophes), l'évolution globale de la composition de la communauté suit ici un patron plus compréhensible et conduit à des différences hautement significatives de répartition d'espèces tolérantes par classe de qualité (sauf entre ME et TME).

En tenant compte du fait que la plupart des 28% (en moyenne) de taxons dans la classe de TBE et la plupart des 17% (en moyenne) de taxons dans la classe de Bon Etat de la Figure 19 ont été en grande partie mal affectés dans cette analyse par rapport à leurs caractéristiques autoécologiques réelles (elles ont toutes été considérées comme des taxons tolérants, ce qui n'est certainement pas le cas vu qu'il y en a une quantité beaucoup plus importante dans les 2 meilleures classes de qualité que dans toutes les autres), la distribution des traits autoécologiques par classe se rapprocherait beaucoup de celle obtenue en 2007 sur la base du jeu de données national mixant tous les types de cours d'eau (Figure 21), passant de sensiblement moins de 50% de taxons tolérants dans la classe de Bon Etat à sensiblement plus de 50% dans la classe d'Etat Moyen, avec des différences significatives entre les deux.

Dans le cadre du présent rapport final, il n'est pas encore possible d'insérer une édition réactualisée et exacte des quelques figures qui précèdent (figures 23 à 25) afin qu'elles utilisent une ventilation auto-écologique biogéographiquement correcte de tous les taxons concernés. En effet, le travail de redéfinition des attributs autoécologiques des taxons de France est actuellement en cours de réalisation, sur la base du jeu de données le plus large dont il était possible de disposer à mi-2018, et ce chantier va encore nécessiter un peu plus d'un semestre avant d'être consolidé.

Malgré l'imprécision d'affectation autoécologique de certains taxons qu'elles incluent, ces figures ont été maintenues car elles présentent cependant l'intérêt d'illustrer et de donner une bonne idée, qui ne pourra que s'améliorer par la suite, de la performance de classification de l'état écologique permise par l'IBD₂₀₀₇, en confirmant bien la **gradation de classification auto-écologique des taxons** en fonction des **classes de qualité attribuées par notre système national d'évaluation diatomique**.

4.3. Liens entre les notes d'IBD₂₀₀₇ et les caractéristiques de saprobie sur une sélection de Très Grands Cours d'Eau

Sur une base de 645 échantillons de diatomées disponibles, concernant uniquement des Très Grands Cours d'Eau (seuil de sélection des bassins versants > 8 000 km²), nous avons réalisé le même type d'analyse que précédemment afin de figurer les caractéristiques des communautés de diatomées vis-à-vis de la **matière organique (altération saprobique)**.

Nous tout d'abord essayé d'étudier la distribution des espèces **sensibles à l'altération organique**, regroupant la classe 1: «Oligosaprobe» et la classe 2: "Bêta-mésosaprobe" de van Dam (1994).

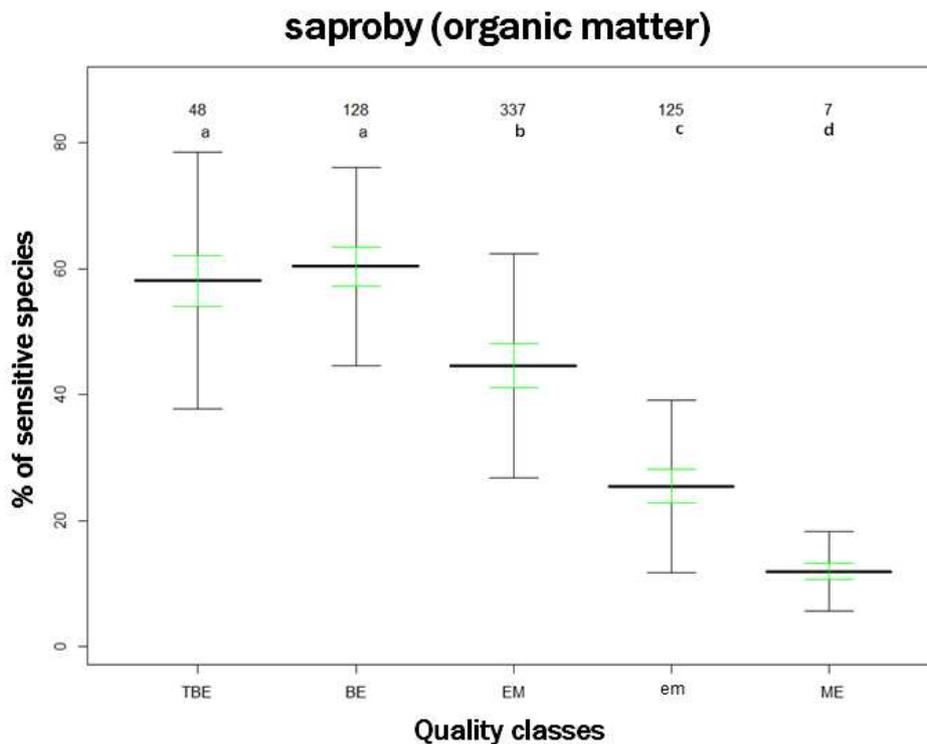


Figure 26 : Pourcentage d'espèces sensibles à l'altération saprobique par classe de qualité dans les TGCE (> 8 000 km²)

Les résultats présentés en **Figure 26** révèlent globalement :

- une absence de différence significative dans la proportion d'espèces sensibles entre la classe TBE et la classe BE, cette catégorie de taxons étant dominante dans la communauté (niveau moyen approchant ou dépassant les 60%),

- une diminution significative des espèces sensibles entre la classe de Bon Etat (BE) et la classe d'Etat Moyen (EM) qui devient dominée par des espèces tolérantes (niveau des espèces sensibles devenant inférieur à 50%, avec une différence significative de distribution entre ces 2 classes),
- Une baisse encore une fois significative et sévère de la proportion de taxons sensibles pour les classes ME et TME.

A nouveau, ces résultats viennent conforter le choix du niveau de seuil G / M qui a été établi pour notre système d'évaluation national.

On peut cependant détecter un fait n'apparaissant pas vraiment logique dans la physionomie générale de distribution des taxons sensibles, en constatant l'inversion de distribution du trait de sensibilité à la saprobie entre la classe de TBE et la classe de BE (même si non significative). Ce résultat suggère l'intervention du même type d'artefact que précédemment pour les attributs trophiques des espèces.

Comme dans le cas de la classification trophique, dans la classification de van Dam, une "Classe 6" de saprobie constituant un fourre-tout regroupait les taxons qui ne pouvaient pas être attribués à une classe saprobique précise, la raison pouvant tenir soit à des données insuffisantes, soit à la quasi-inexistence de l'espèce dans les comptages néerlandais. Cela conduit donc à un déficit d'information écologique pour ces taxons pour une raison biogéographique.

Nous avons donc essayé de diagnostiquer la situation en étudiant la répartition de ces taxons inconnus par classe de qualité (cf. **Figure 27**).

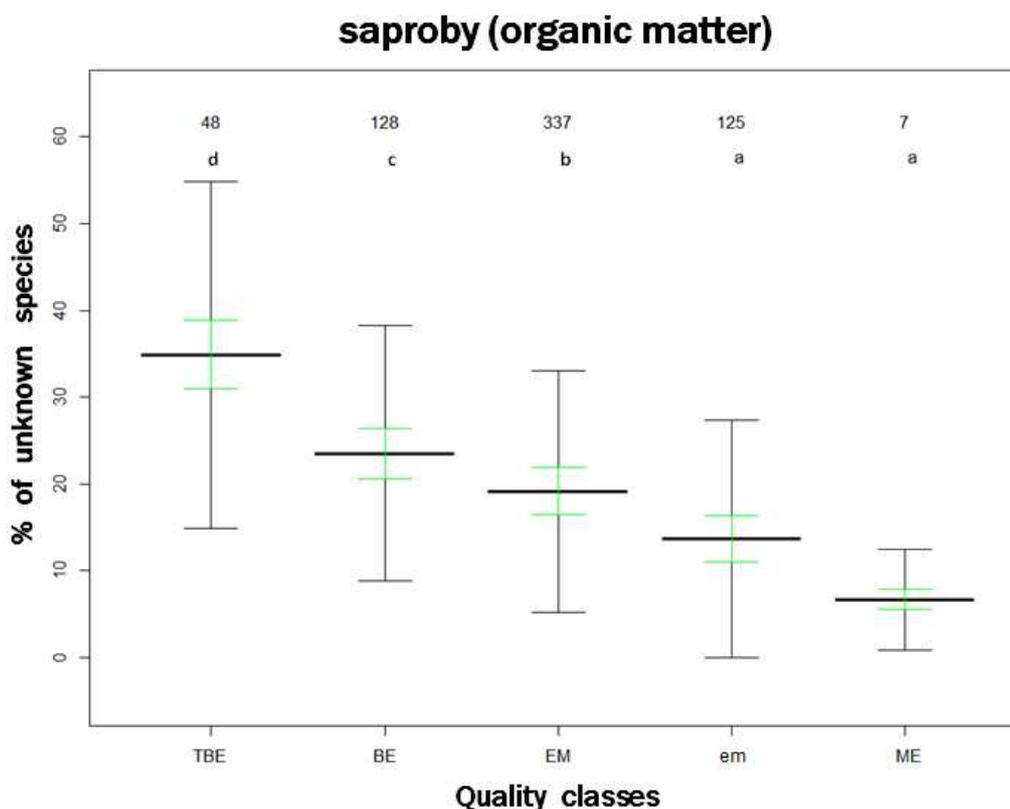


Figure 27: Distribution des espèces au statut indéterminé vis-à-vis de la saprobie au long du gradient des 5 classes de qualité

Comme pour la trophie, une part assez importante de taxons dans les communautés de diatomées des TGCE ont, selon van Dam, une classification inconnue ou mal définie vis-à-vis de la saprobie et de plus, leur distribution n'est pas équilibrée entre les classes, mais très sensiblement plus forte en classe de TBE (> 35% en moyenne) et en BE (> 23% en moyenne), que dans les classes de qualité plus dégradée. Le fait que ce positionnement est orienté vers les 2 meilleures classes de qualité des cours d'eau de notre pays (concernant en particulier la Dordogne et la Durance) suggère donc qu'un pourcentage dominant de ces taxons mal décrits aux Pays-Bas est doté d'une écologie adaptée aux milieux naturels de bonne à très bonne qualité écologique, et est en fait **sensible à l'altération saprobique**.

Ce gradient d'espèces indéterminées fonctionne donc à l'inverse de la pente de répartition des espèces sensibles et il vient atténuer et flouter très sensiblement les résultats présentés en Figure 26, avec pour effet de diminuer artificiellement le niveau de "taxons saprophyte sensibles", dans les classes de TBE et de BE. L'anomalie conduisant à une inversion de classification entre classe de TBE et de BE est certainement aussi en rapport avec cette lacune de connaissance autoécologique de certains taxons diatomiques dans les conditions des Pays-Bas, qui ne possèdent pas de rivières de type TGCE arrivant directement de massifs montagneux de haute altitude (du type Dordogne ou Durance).

À l'inverse, la qualité écologique des cours d'eau de plaine des Pays-Bas étant souvent de niveau moyen, voire dégradé, van Dam a été en mesure de décrire et de qualifier l'autoécologie saprobique de la plupart des taxons tolérants que nous trouvons aussi en France, illustrant bien le fait que cette catégorie d'espèces a été plus facilement trouvée et décrite dans les hydrosystèmes des Pays-Bas.

De même que pour les orthophosphates (phosphore soluble et biodisponible), nous avons alors essayé d'exprimer la distribution de l'autécologie vis-à-vis de la saprobie en montrant l'augmentation des taxons tolérants au fur et à mesure de la dégradation des classes de qualité. Les résultats de ce mode inverse de construction sont présentés dans la **Figure 28**.

Là aussi, cette façon de faire s'accommodant d'une certaine proportion d'informations autoécologiques erronées pour notre contexte national ne prétend pas pour l'instant donner un résultat autoécologique exact. Cependant, elle peut préfigurer la physionomie qu'une métrique basée sur l'apparition de taxons tolérants à l'altération organique pourrait prendre, une fois que le travail de révision des caractéristiques autoécologiques des taxons Français sera terminé (travail actuellement en cours, dont le rapport final devrait être rendu disponible vers la mi-2019).

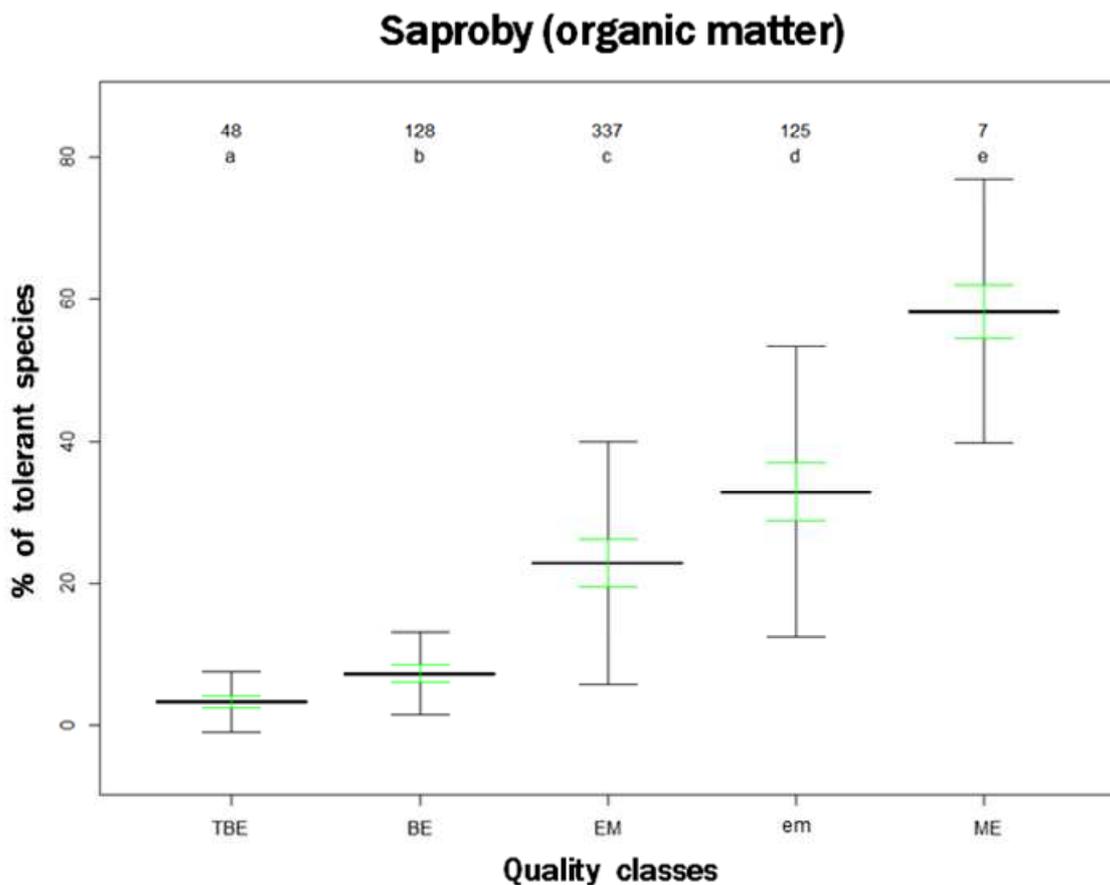


Figure 28: Pourcentage d'espèces tolérantes à la saprobie par classe de qualité dans les TGCE (> 8 000 km²)

Les résultats sont assez analogues à ceux concernant le phosphore dissous :

- La différence entre la classe de TBE et celle de BE est assez modeste, mais apparaît significative à l'analyse.

- La différence entre la classe de BE et celle d'Etat Moyen (EM) est également significative, mais sensiblement plus importante (on passe d'une moyenne de 6% à une moyenne de près de 25% d'espèces tolérantes), révélant une modification sensible des communautés de diatomées avec l'augmentation de ce genre de pression.

Ici encore, les résultats présentés dans la Figure 28 viennent conforter, par un autre type d'approche basée sur l'autoécologie des taxons, la pertinence du calage de basculement entre classe de Bon Etat et Classe d'Etat Moyen dans le système d'évaluation écologique proposé par l'Etat Français, en conformité avec la définition normative du bon état contenue dans la DCE.

4.4. Conclusions sur les relations pression-impact de l'IBD₂₀₀₇ sur TGCE

Dans le cadre de l'étude nationale TGCE et aussi, en vue de répondre à des demandes formulées par le GIG « Large Rivers », il a été procédé, à chaque fois ou c'était possible, à une mise en comparaison d'analyses de données TGCE issues de requêtes PANDORE 2015 et 2017 avec les analyses antérieures effectuées en 2007 sur un jeu de données national complet (incluant tous les types et toutes les tailles de cours d'eau), en vue de venir à l'appui de l'illustration de performance du prototype de système d'évaluation TGCE basé sur l'IBD₂₀₀₇.

De façon non surprenante, le test des relations pression-impact réalisé sur une sélection de relevés uniquement issus de TGCE a conduit à la **diminution du gradient interne de pressions anthropiques représenté** (absence de vrais sites de référence, diminution de représentation de sites très altérés), modifiant et venant atténuer le niveau global des relations pression-impact antérieurement établies entre différents gradients de qualité abiotique et la réponse de notre système national d'évaluation basé sur l'IBD₂₀₀₇, tous cours d'eau confondus.

Dans ce contexte général de réduction d'amplitude globale des gradients d'anthropisation représentés, un autre facteur suspect d'avoir contribué de façon encore plus importante au floutage des relations pression-impact construites sur TGCE, par rapport à des analyses tous cours d'eau confondus, résulte du fait que l'altération de l'état écologique est le plus souvent occasionnée par un contexte multi-pressions (ex : aval de rejet de STEP). Or, dans le cadre de cet exercice d'intercalibration, **l'unique pression** ayant servi de base à l'inter-comparaison des systèmes d'évaluation des Etats-Membres était la **trophie** (concentrations en PO₄). En revanche, l'indice diatomique sur lequel se base tout le dispositif d'évaluation d'état écologique des cours d'eau de France est l'IBD₂₀₀₇, dont la formulation a été basée sur la prise en compte des gradients présentés par **7 descripteurs différents d'anthropisation**. Par construction, cet indice est donc révélateur d'altérations de l'état écologique occasionnées par des pressions anthropiques diverses entraînant des conséquences sur la composition physico-chimique et le niveau d'enrichissement en nutriments des cours d'eaux, et sa réponse intégrative mesure en fait un niveau global d'altération lié à un contexte multi-pressions.

Pour les 2 grands types de raisons précitées, Il est donc normal que le coefficient de corrélation obtenu sur un jeu de données TGCE vis-à-vis d'une mono-pression anthropique (soit ici PO₄, responsable de l'eutrophisation des cours d'eau) conduise à l'expression d'une relation non-optimale et bruitée, ce qui a été le cas pour notre jeu de données Français. Cependant, l'aspect modeste des relations pression-impact montrées dans notre contexte national n'a pas généré de difficultés particulières, vu que la même problématique se posait de façon générale sur tous les TGCE européens regroupés dans le jeu de données d'intercalibration (là aussi, niveau très modeste de la relation globale obtenue) et qu'il était même très difficile voire impossible à certains pays d'en illustrer une (cas par exemple de la Roumanie, de la Bulgarie) faute d'un gradient suffisamment différencié d'altération au sein d'un même et parfois unique hydrosystème fluvial (ex : le Danube).

Quoi qu'il en soit, les résultats des analyses réalisées et présentées aux 2 chapitres qui précèdent :

- montrent qu'il est plus difficile d'illustrer des relations pression-impact de bonne qualité sur des TGCE que sur un assortiment de cours d'eau tout-venant, incluant une large majorité d'hydrosystèmes de taille plus réduite,
- confirment que l'investigation des relations pression-impact sur une base moyenne annuelle n'est pas la façon la plus adéquate de procéder, au moins dans le cas du compartiment des diatomées benthiques et pour des descripteurs plutôt révélateurs d'un niveau d'altération anthropique. Dans ce type de cas, une base d'intégration saisonnière calée sur une durée de 3 mois (ordre indicatif de durée de vie moyenne d'un biofilm) donne des résultats plus convenables,
- n'ont pas vraiment permis de montrer de façon convaincante qu'à cette échelle de taille de cours d'eau, la réponse de l'IBD 2007 est plus représentative d'une réponse intégrée à un gradient d'altération composite qu'à des pressions simples (ce qui serait attendu vu la façon dont il a été conçu et calé...),

- ont de ce fait permis de révéler que les stratégies ponctuelles de prélèvements d'eau donnent une image défectueuse (encore moins représentative que pour des cours d'eau plus petits, et la plupart du temps sous-estimée) de l'exposition chimique aux nutriments réellement intervenue *in situ*. La réponse biologique restituée par l'IBD apparaît en effet nettement plus robuste et intégrative du temps pour rendre compte de l'exposition réelle à laquelle ont été soumises les diatomées (venant ici en confirmation du précepte prôné par la DCE, qui a mis en avant l'intérêt d'utiliser des maillons-clés biologiques),
- soulignent que les traits autécologiques compilés par van Dam fournissent une bonne base de travail et produisent des résultats plutôt consistants et robustes ; mais aussi qu'ils nécessitent une adaptation biogéographique et une extension vers une liste plus complète des espèces de France en vue de procurer des résultats optimaux vis-à-vis de notre territoire,
- illustrent un ajustement plutôt satisfaisant et justifié du calage de seuil entre BE et EM de notre dispositif national envers plusieurs critères de pression et/ou de réponse biologique des communautés diatomiques,

Même dans un contexte de référentiel biogéographiquement décentré par rapport aux conditions de nos hydrosystèmes naturels, l'avant-dernier item de la liste précédente, basé sur des attributs autoécologiques des communautés diatomiques, produit des résultats plus robustes et moins erratiques que les résultats de concentrations intégrées en nutriments basés sur des échantillons ponctuels, confirmant par un autre type d'approche la qualité de la bio-intégration réalisée par ce maillon biologique. Un travail est donc actuellement en cours pour formaliser différents attributs autoécologiques des espèces diatomiques rencontrées et positionnées dans notre contexte biogéographique national. Outre l'intérêt en soi de cette formalisation, les informations produites seront aussi utilisées pour fonder diverses métriques de diagnostic notamment utilisables vis-à-vis des pollutions par les nutriments et, sur un plan plus général, de certaines altérations physico-chimiques des eaux émanant d'infrastructures industrielles publiques (STEPS) ou privées (usines chimiques, industries agro-alimentaires etc...).

5. PROPOSITION ET STABILISATION D'UN SYSTEME FRANCAIS D'EVALUATION DES TGCE

5.1. Eléments généraux et de contexte

Dans le cadre de la mise en application de la DCE en France métropolitaine, du fait de certaines difficultés spécifiques dont l'absence quasi-totale de situations de référence vraie sur tout notre territoire, la mise au point d'un système d'évaluation diatomique dédié aux très grands cours d'eau (TGCE) n'a été entreprise qu'assez tardivement, une fois que les systèmes d'évaluation dédiés aux cours d'eau plus petits, plus faciles à élaborer dans le respect des recommandations des documents-guides ECOSTAT et REFCOND, étaient mis en place et avaient déjà fait l'objet d'une intercalibration européenne.

A partir de son entrée en vigueur en Octobre 2000, la DCE a établi et mis en place un nouveau cadre conceptuel qui préside à l'objectif et à la façon d'évaluer plus judicieusement les écosystèmes aquatiques. Elle considère notamment qu'un milieu aquatique doit se révéler convenable, non uniquement pour les usages sociétaux qu'il porte, mais aussi et avant tout pour les compartiments biologiques qu'il héberge, qui doivent pouvoir y boucler normalement et sans problème particulier toutes les phases de leur cycle de vie. Cette Directive place donc les compartiments biologiques au centre du dispositif d'évaluation et définit un certain nombre de maillons-clés bio-intégrateurs jugés pertinents selon le type de masse d'eau à évaluer.

Parmi un ensemble de recommandations destinées à faire décliner dans la pratique cette nouvelle philosophie, la DCE a notamment demandé d'évaluer le **degré d'altération de l'état écologique** des hydrosystèmes soumis à des pressions anthropiques par **écart au niveau de référence adéquat**.

Cette construction méthodologique impose d'une part d'avoir repéré la trame naturelle permettant de définir des **zones naturelles homogènes** sur lesquelles il est possible d'appliquer le même système d'évaluation (ces zones homogènes, appelées hydroécorégions ou HERs, ayant le plus souvent fait l'objet d'un zonage géographique national basé sur différentes catégories de variables abiotiques influentes sur les caractéristiques naturelles des habitats). Elle implique d'autre part que la mesure du **niveau d'altération** subi sur un site donné se fasse par écart avec le **niveau de référence adéquat** donné par la méthode nationale sur cette zone. En pratique, ce niveau de référence découle d'un calcul de moyenne établi sur un assortiment de sites de référence internes à la zone considérée et jugés représentatifs de ses conditions naturelles, hors contexte mesurable de dégradation anthropique.

Pour les hydrosystèmes de petite, moyenne et grande taille, dans la mesure où il existait des situations de référence valides pour la zone naturelle concernée ou pour d'autres zones naturelles aux caractéristiques très comparables, l'application de ce principe de construction n'a pas posé problème, ou n'a nécessité que quelques aménagements à la marge pour conduire à la définition d'un cadre d'évaluation convenable. C'est pourquoi des dispositifs nationaux d'évaluation ont pu être assez rapidement proposés, puis intercalibrés.

Cependant, par rapport aux cours d'eau plus petits, l'évaluation des TGCE de France, mais aussi des autres Pays-Membres, a posé une difficulté méthodologique supplémentaire, à savoir qu'il n'existe plus ou pratiquement plus, à l'échelle au moins de toute l'Europe de l'Ouest, de sites de référence valides sur des cours d'eau de cette échelle de taille, inévitablement soumis à des pressions anthropiques plus ou moins intenses, mais de toute façon omniprésentes.

Dans un tel contexte, l'évaluation des très grands cours d'eau va donc devoir se baser sur un **niveau de référence par défaut** (= *least disturbed condition*). Cette notion un peu vague et arbitraire est difficile à travailler à l'échelle d'un simple pays, d'autant que le niveau de référence par défaut qui serait choisi par différents pays et sur différents types de cours d'eau a toutes les chances de ne pas être homogène à l'échelle de l'Europe et de pouvoir être contesté, à la hausse comme à la baisse, par d'autres pays-Membres situés dans un contexte naturel, sociétal et économique différent.

Pour ce chantier, plutôt que de se baser uniquement sur une étude nationale, dont l'utilité est bien sûr incontestable, mais ne suffit pas pour garantir l'obtention à terme d'un visa de DCE-compatibilité, il valait mieux essayer de s'inscrire, selon une opportunité de circonstance, dans la réalisation d'un exercice européen d'intercalibration nouvellement programmé (sachant qu'il en avait déjà été réalisé un en 2011-2012, coordonné par le X-GIG « Large Rivers»). Une telle opportunité permet d'aborder dans un cadre collectif la fixation d'un niveau de référence par défaut, défini autant que possible sur une base consensuelle par type européen de Très Grand Cours d'Eau, mais aussi d'obtenir des repères communs avec d'autres pays-Membres pour une déclinaison en classes de qualité, permettant au final le calage d'un système national dans un cadre garantissant d'emblée une évaluation plus cohérente d'un Pays-Membre à l'autre.

Dans les faits, même si l'intention initiale était de participer dès que possible à un nouvel exercice d'intercalibration TGCE à partir du moment où une telle opportunité se présenterait, l'adhésion à une telle opportunité était programmée dans le cadre de l'actuelle Fiche-Action Irstea-AFB « Amélioration des méthodes végétales d'évaluation en cours d'eau » 2016-2018, et pas de la fiche précédente 2013-2015.

Or l'organisation de cet exercice collectif, qui constituait par ailleurs une opportunité à saisir, s'est mise en place sans anticipation (réception d'un premier message circulaire à la mi-Juin 2015, sans qu'aucune annonce préalable d'organisation n'ait été circulée par l'UE un peu à l'avance) et a demandé une réactivité sans faille, en période estivale et à marche forcée, afin de collecter, formater et envoyer dans les temps les éléments et jeux de données demandés pour pouvoir participer effectivement à l'exercice (deadline en septembre 2015).

L'inscription un peu impromptue à cet exercice a de fait bousculé les priorités de travail du moment et nécessité une ré-allocation de force de travail (permanents, CDD) consacrée en toute priorité à ce chantier, s'inscrivant malheureusement en compétition avec d'autres actions prévues dans le cadre de la même fiche-action, venant en désorganiser l'avancement et retarder la fourniture de certains autres livrables prévus (retards encore en cours de résorption actuellement). En effet, pour pouvoir respecter la deadline imposée pour une participation nationale effective, il a fallu très rapidement procéder à des investigations de données et fournir un jeu de données national comportant une sélection parmi les données existantes, la contrainte étant que ce jeu de données soumis à l'intercalibration UE reste représentatif de nos très grands cours d'eau et des gradients d'altération présents au niveau de la France. Cet état de fait a nécessité un travail approfondi et une prise de recul rapide sur les référentiels nationaux, et a aussi exigé un travail important de post-traitement de données géochimiques-hydrochimiques et géographiques afin de pouvoir typifier nos hydrosystèmes par rapport à des descripteurs typologiques du GIG non-utilisés au niveau national, donc présentant des difficultés de disponibilité.

Suite à ce travail estival, un premier jeu de données-test a pu être fourni le 15 Septembre 2015, accompagné de différents tableaux requis sur les sites, les relevés diatomiques et les données abiotiques collectées, ainsi que de notes techniques en langue anglaise permettant de décrire d'une part notre méthode indicelle nationale (IBD₂₀₀₇) et ses performances en termes de relations pression-impact, et d'autre part notre système national d'évaluation proposé (qui n'avait pas encore d'existence réelle à cette date). Un jeu de données national officiel consolidé (sélection de 221 relevés acquis sur 99 sites différents), accompagné d'un complément d'informations descriptives complémentaires requises, a fait l'objet d'un 2^{ème} et dernier envoi en date du 15-12-2015.

La progression de notre prototype de dispositif national d'évaluation TGCE, en lien étroit avec le fonctionnement de l'exercice d'intercalibration et ses résultats d'étape, s'est ensuite faite au moyen d'échanges et d'itérations successives intervenues entre septembre 2015 et la mi-Juin 2016.

Le constat de résultats n'entrant pas dans les biais requis a ainsi conduit à plusieurs évolutions-ré-adaptations du prototype national de dispositif proposé pour la France, jusqu'aux derniers ajustements nationaux soumis en date des 16 et 17 Juin 2016 (deadline imposée par les organisateurs pour la fin d'exercice), après tenue d'une visioconférence de concertation finale avec la Direction de l'EAU et l'AFB/DAST.

Au final, l'exercice, qui a été considéré comme clos et agréé UE en Février 2018, a nécessité en tout une année complète (15 Juin 2015 - 21 Juin 2016) de contacts de travail très serrés incluant propositions ou fourniture Française de divers éléments dont jeu de données national initial ou amendé (cf. changements de typologie et de niveaux d'évaluation) – tests du GIG – réactions – envoi de questions ou remarques sur le plan méthodologique - nouvelles itérations, jusqu'à la version finale de dispositif jugée adéquate au sein de l'exercice. Ces échanges nécessitant une réactivité très importante au fil de l'eau ont atteint une capacité de travail de 1,6 à 1,7 plein-temps sur un an et ont représenté sensiblement plus de 2 pleins-temps en cumulé sur la durée, dépassant largement le volume de travail initialement prévu et financé sur ce chantier et empiétant sur les moyens prévus pour d'autres types actions.

Par rapport à la proposition initiale, les différentes versions modificatives de notre dispositif national ont à chaque fois été proposées sur sollicitation du GIG, en réaction par rapport à l'observation de résultats décalés dans le cadre de l'exercice d'intercalibration UE, qui auraient conduit à une intercalibration « non OK » pour la France. Tout en essayant de tenir compte de spécificités évidentes de certains de nos TGCE de France, qui semblent avoir très peu d'équivalent en Europe, mis à part sur de grands hydrosystèmes Italiens situés juste à l'aval de massifs montagneux, elles ont visé à ramener progressivement notre dispositif national vers une vision de l'état écologique plus partagée au niveau des Etats-Membres participants.

Du fait de ce mode de fonctionnement par itérations successives sur le mode « intercomparaison UE ⇒ réadaptation nationale », et sans entrer dans tous les détails de la procédure suivie, la structuration de la suite du présent chapitre relatif à la proposition et à l'évolution de de notre système national a forcément été présentée de façon un peu linéaire et descriptive, se basant sur les différentes itérations et boucles de rétro-actions successivement réalisées.

5.2. Dispositif d'évaluation en vigueur, répercussions vis-à-vis des TGCE

Sur le plan général de l'évaluation des cours d'eau de France, des grilles nationales d'évaluation diatomique basées sur les notes d'IBD et déclinées en EQRs à partir des niveaux moyens de référence atteints par région naturelle « diatomées » (regroupements d'HER1) servent au calcul d'état écologique dans le cadre des réseaux, ainsi qu'au rapportage régulier d'état écologique auprès de l'Union Européenne prévu par la DCE.

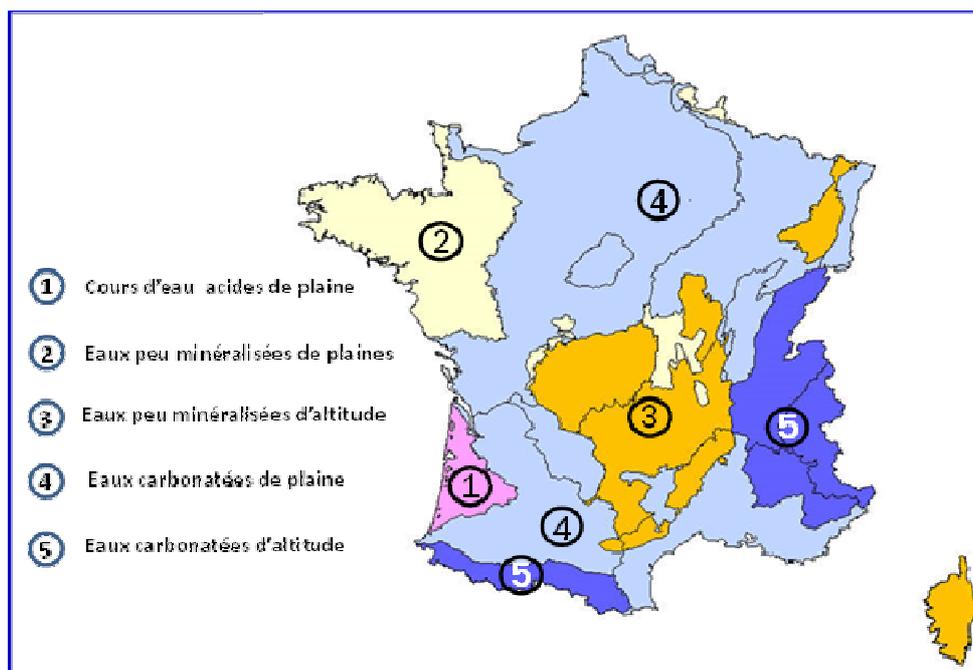


Figure 29 : Zonage des régions diatomiques naturelles en vigueur pour l'évaluation d'état écologique (Version 2 de 2012 incluse dans l'Arrêté Evaluation du 27 Juillet 2015)

Le zonage naturel de référence actuellement utilisé pour l'application de ces grilles est présenté en **Figure 29**. Il a été obtenu suite à l'établissement d'une biotypologie des flores diatomiques de référence utilisant un réseau neuronal non supervisé (SOM ou Self-Organizing Map).

Les résultats de la SOM ont permis d'aboutir à 5 régions « Diatomées » présentant des types de flores naturelles différentes, les contours géographiques de ces zones diatomées correspondant à des regroupements d'HER de niveau 1 définies au préalable au niveau national (**REF. 10** : Wasson et al, 2004). En fonction d'une actualisation d'analyse biotypologique réalisée par Bottin et al (**REF. 11** : Rosebery et al, 2011), ce zonage a un peu évolué par rapport au PGME précédent, du fait d'une augmentation du jeu de données disponibles et de l'utilisation en cumul du contenu de 3 relevés annuels par site de référence (au lieu de relevés individuels annuels pour la version précédente) pour typifier les flores tout en tamponnant l'effet de variations annuelles.

Les grilles actuelles utilisées pour l'évaluation de l'état écologique à partir des diatomées benthiques des cours d'eau sont présentées en **Figure 30** ci-dessous. Elles servent à évaluer les cours d'eau de France de toute taille, mais n'ont pas été spécifiquement travaillées et raisonnées pour une bonne prise en compte des spécificités des TGCE. Les 5 grilles du haut de la figure sont données en valeurs approximées d'IBD₂₀₀₇ (arrondies au 1/10^{ème} de point). La grille d'évaluation du bas de la figure, exprimée en valeurs exactes d'EQR (ou équivalents de qualité écologique), est celle qui fait référence pour l'évaluation et l'intercalibration.

Pour intégrer les spécificités propres à chacun des 5 ensembles naturels concernés, elle re-norme l'échelle d'évaluation entre le niveau médian de référence pour la zone naturelle considérée (EQR = 1) et la note indiciaire minimale atteignable dans la zone considérée (EQR = 0). Elle est unique pour les 5 types de cours d'eau.

IBD ₂₀₀₇	REF	H/G	G/M	M/P	P/B	MIN
Type 1	18,4	17,596	15,452	12,370	9,020	5,0
Type 2	17,4	16,416	13,792	10,020	5,920	1,0
Type 3	19,0	18,160	15,920	12,700	9,200	5,0
Type 4	18,1	17,074	14,338	10,405	6,130	1,0
Type 5	20,0	19,100	16,700	13,250	9,500	5,0

Biological Quality Element	Biological Index	Reference (median)	Ecological status class limits (in IBD-EQRs)				
			H/G	G/M	M/P	P/B	MIN
Phytobenthos (diatoms)	IBD ₂₀₀₇	1	0,94	0,78	0,55	0,30	0

Figure 30 : Grilles d'évaluation diatomique actuelles (Arrêté Evaluation du 27-07-2015)

Faute d'analyse de données assez solide sur laquelle s'appuyer, une grille d'évaluation des très Grands Cours d'Eau de France encore sommaire, largement basée sur un jugement-expert et contenant un nombre en réduction, selon les versions, de cases grisées avec dièse (cours d'eau non évaluables dans l'état actuel des données disponibles et de nos connaissances), a été publiée à 2 reprises :

- 1) Arrêté d'Evaluation du 25 Janvier 2010, très incomplet sur cette typologie de cours d'eau ;
- 2) Plus récemment, Arrêté Evaluation du 27 Juillet 2015, dont le contenu était destiné à rester en vigueur jusqu'à la fin 2018 (cette échéance correspondant à la fin du demi-PGME en cours). Dans cette 2^{ème} version d'Arrêté, le contenu a été complété sur certains types de TGCE, après examen-expert et le plus souvent, par simple propagation vers l'aval du niveau d'évaluation appliqué immédiatement à l'amont.

La grille d'application de l'Arrêté d'Evaluation du 27-07-2015 présentée en **Figure 31** page suivante, qui a été en vigueur jusqu'au 31 Décembre 2018, a été produite sur une base d'expertise et en partant du présupposé que l'échelle de taille des cours d'eau n'influence pas, ou influe vraiment à la marge, par rapport à des cours d'eau plus petits, sur le niveau naturel de notation par le maillon diatomique.

		Valeurs inférieures des limites de Classes d'Etat Ecologique par type (IBD 2007)					
		Rangs (bassin Loire-Bretagne)	8, 7	6	5	4	3, 2, 1
IBD 2007		Rangs (autres bassins)	8, 7, 6	5	4	3	2, 1
Hydrocôrégions de niveau 1		Cas général, cours d'eau exogène de l'HER de niveau 1 indiquée ou HER de niveau 2	Très Grands	Grands	Moyens	Petits	Très Petits
20	DEPOTS ARGILLO SABLEUX	Cas général		17-14,5-10,5-6		17-14,5-10,5-6	17-14,5-10,5-6
		HER2 n°31, 33, 39, 119 (Artois-Picardie)		17-14,5-10,5-6		17-14,5-10,5-6	
		Exogène de l'HER 9		17-14,5-10,5-6			
21	MASSIF CENTRAL NORD	Exogène de l'HER 21		18-16-13-9,5	18-16-13-9,5		
		Cas général		18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5
3	MASSIF CENTRAL SUD	Cas général		18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5
		Exogène de l'HER 19		#			
		Exogène de l'HER 8		#			
17	DEPRESSIONS SEDIMENTAIRES	Exogène de l'HER 19 ou 8	17-14,5-10,5-6				
		Cas général		16,5-14-10,6	16,5-14-10,6	16,5-14-10,6	16,5-14-10,6
		Exogène de l'HER 3 ou 21	#	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5
15	PLAINE SAONE	Exogène de l'HER 5		19-16,5-13-9,5	19-16,5-13-9,5		
		Cas général	17-14,5-10,5-6		17-14,5-10,5-6		17-14,5-10,5-6
		Exogène de l'HER 10	17-14,5-10,5-6				
5	JURA / PRE-ALPES DU NORD	Cas général		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
		Exogène de l'HER 2	19-16,5-13-9,5		19-16,5-13-9,5		
TTGA	FLEUVES ALPINS	Cas général	#				
		Exogène de l'HER 3 ou 21		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
2	ALPES INTERNES	Cas général		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
		Exogène de l'HER 3 ou 8		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
7	PRE-ALPES DU SUD	Cas général		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
		Exogène de l'HER 2	17-14,5-10,5-6	19-16,5-13-9,5			
		Exogène de l'HER 2 ou 7		19-16,5-13-9,5			
6	MEDITERRANEE	Exogène de l'HER 7		19-16,5-13-9,5			
		Exogène de l'HER 8	17-14,5-10,5-6	18-16-13-9,5			
		Exogène de l'HER 1		19-16,5-13-9,5			
		Cas général	17-14,5-10,5-6		17-14,5-10,5-6		17-14,5-10,5-6
8	CEVENNES	Cas général		18-16-13-9,5		18-16-13-9,5	
		A-her2 n°70		18-16-13-9,5		18-16-13-9,5	
16	CORSE	A-her2 n°22		18-16-13-9,5		18-16-13-9,5	
		B-her2 n°88	18-16-13-9,5		18-16-13-9,5		18-16-13-9,5
19	GRANDS CAUSSES	Cas général				17-14,5-10,5-6	
		Exogène de l'HER 8		18-16-13-9,5			
11	CAUSSES AQUITAINS	Cas général				17-14,5-10,5-6	17-14,5-10,5-6
		Exogène de l'HER 3 et/ou 21	17-14,5-10,5-6	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	
14	COTEAUX AQUITAINS	Exogène des HER 3, 8, 11 ou 19	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6		
		Exogène de l'HER 3 ou 8		17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6
		Cas général		17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6
13	LANDES	Exogène de l'HER 1	17-14,5-10,5-6	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	
		Cas général		17,5-15,5-12,5-9	17,5-15,5-12,5-9	17,5-15,5-12,5-9	17,5-15,5-12,5-9
1	PYRENEES	Cas général		19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5	19-16,5-13-9,5
		A-Centre-Sud		16,5-14-10,6	16,5-14-10,6	16,5-14-10,6	16,5-14-10,6
12	ARMORICAIN	B-Ouest-Nord-Est		16,5-14-10,6	16,5-14-10,6	16,5-14-10,6	16,5-14-10,6
		Cas général	17-14,5-10,5-6		17-14,5-10,5-6		17-14,5-10,5-6
TTGL	LA LOIRE	Cas général	17-14,5-10,5-6				
		A-her2 n°57			17-14,5-10,5-6	17-14,5-10,5-6	
9	TABLES CALCAIRES	Cas général	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6
		Exogène de l'HER 10		17-14,5-10,5-6	17-14,5-10,5-6		
		Exogène de l'HER 21	17-14,5-10,5-6	18-16-13-9,5	18-16-13-9,5		
10	COTES CALCAIRES EST	Exogène de l'HER 21		18-16-13-9,5	18-16-13-9,5		
		Cas général	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6	17-14,5-10,5-6
		Exogène de l'HER 4		18-16-13-9,5	18-16-13-9,5		
4	VOSGES	Cas général		18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	18-16-13-9,5
		Exogène de l'HER 10	16,5-14-10,6				
22	ARDENNES	Cas général		16,5-14-10,6	16,5-14-10,6	16,5-14-10,6	16,5-14-10,6
		Exogène de l'HER 4			17-14,5-10,5-6		17-14,5-10,5-6
18	ALSACE	Cas général			17-14,5-10,5-6		17-14,5-10,5-6
		Exogène de l'HER 4		18-16-13-9,5	18-16-13-9,5	18-16-13-9,5	

a=b-c-d : a = limite inférieure du très bon état, b = limite inférieure du bon état, c = limite inférieure de l'état moyen, d = limite inférieure de l'état médiocre
 # : absence de référence. En gris : type inexistant

Figure 31 : Grille d'évaluation diatomique des TGCE actuellement en vigueur ('Arrêté du 27-07-2015)

Avec la prise de recul permise par l'élargissement du référentiel national disponible, il est devenu de plus en plus évident que, si cette atteinte des plus hautes notes est possible à **certaines dates** sur certains grands hydrosystèmes de bonne à très bonne qualité, elle est nettement moins systématique sur une série de plusieurs relevés prélevés dans un tel site, qui intègrent une part de variabilité interannuelle.

En effet, même en l'absence d'anthropisation notoire, les types de forçages s'appliquant sur de grands hydrosystèmes provoquent une sélection de flores typiques de parties plus aval de cours d'eau, entraînant une baisse au moins statistique de notes d'indice parfois simplement liée à un type de fonctionnement physique différent des cours d'eau. Le type de minéralisation naturelle de l'eau peut aussi exercer sa part d'influence sur le niveau de notation indicielle.

La pratique initiale « faute de mieux » a permis de pallier momentanément l'absence de méthodologie facilement applicable, en l'absence de stations de référence vraie jugées valides et tant que les jeux de données nationaux n'étaient pas suffisants pour permettre une analyse de données consistante.

A l'appui de cette pratique antérieure, il s'est vérifié qu'un grand à très grand cours d'eau alpin ou montagnard avec une partie importante de son bassin versant amont située en altitude (type Durance, Dordogne amont etc...) peut procurer, en l'absence de problème notable d'anthropisation et au moins à certaines dates, des niveaux de notation indicielle aussi hauts, ou pratiquement aussi hauts que dans des cours d'eau de plus petite taille au statut de référence, ou peu anthropisés, placés dans des amonts de bassins versants et de statut naturellement oligotrophe voire mésotrophe, comme il peut en exister dans des environnements de plus faible altitude.

Mais bien sûr, le postulat de départ (atteinte possible des plus hautes notes même sur certains TGCE préservés) était certainement un peu simplificateur ou réducteur par rapport à la réalité du fonctionnement de ces très grands systèmes, qui soumettent forcément les cortèges biologiques qu'ils hébergent à des forçages dont certains (physique des écoulements, profondeur, température, oxygénation de la colonne d'eau, nature et granulométrie des substrats sédimentaires présents sur le fond du lit etc...) sont dépendants de la taille de l'hydrosystème et du positionnement du site suivi dans le continuum longitudinal amont-aval (cf. [REF. 9](#)).

Dans le dispositif d'évaluation actuellement en vigueur, l'affectation initiale des grilles de notation appliquées dans l'ellipse violette de la **Figure 31** l'a été à l'époque sans étude spécifique basée sur l'analyse en propre d'un jeu de données TGCE, donc par propagation amont-aval du système d'évaluation sous contrôle-expert.

Lorsque l'évaluation sur une telle base (certainement perfectible...) était considérée comme applicable sur ce type de cours d'eau du point de vue des experts-maillon, les valeurs contenues dans ces cases correspondent à une propagation à l'identique vers l'aval de la grille d'évaluation (basée sur des EQRs) utilisée sur le même cours d'eau d'un ordre de taille plus faible (cours d'eau de moyens à grands) dans le même ensemble géographique (HER de niveau 1). Les cases comportant un dièse (cas particuliers de la Loire et de l'Allier dans l'HER 17, du Rhône et du Rhin, qualifiés de TTGA dans la typologie nationale) n'ont pas été jugées évaluables de façon raisonnable selon ce principe.

Cependant, certaines situations particulières pour lesquelles cette façon de procéder est un peu simpliste ont pu être repérées :

- Notion de grand cours d'eau traversant une HER donnée alors qu'il provient de façon dominante d'autres contextes géo-climatiques (cas un peu délicat de la Loire et de l'Allier, déjà évoqué, cas de la Meuse etc...) ;
- Problèmes de confluence de cours d'eau de typologie très différente, notamment sur le plan physique et hydrochimique, en fonction de l'implantation géo-climatique de leur bassin versant, et dispositif d'évaluation à adopter à l'aval de cette confluence (ex : le Rhône à partir de sa confluence avec la Saône à Lyon, qui ne peut plus être judicieusement évalué sur la même base de Très Grand Cours d'Eau Alpin). Outre cet exemple très illustratif, d'autres cas d'espèce un peu différents, mais nécessitant un raisonnement et un traitement particulier, ont aussi été repérés ici ou là.

Avec le temps, il est donc devenu de plus en plus évident que les TGCE doivent être évalués sur la base de grilles particulières tenant compte du positionnement des sites concernés sur secteurs aval des cours d'eau, l'une des difficultés méthodologiques à résoudre étant qu'ils ne présentent plus aucune station de référence valide (d'où la nécessité d'utiliser un niveau de référence par défaut, et la difficulté particulière liée à la fixation de son niveau).

De ce fait, la mise en place de la présente étude nationale avait pour but de réviser ce dispositif national d'évaluation TGCE pré-existant, très perfectible :

- en s'appuyant sur le maximum de données nationales actuellement disponibles (les possibilités d'analyse d'un corpus suffisant de données sur cette typologie de taille de cours d'eau étant difficilement envisageables jusqu'à un passé assez récent) ;
- en tenant compte de la pratique et du calage de niveau de nos pays voisins. Dans ce contexte, la participation à l'exercice collectif d'intercalibration « Large Rivers » réalisé entre la mi-2015 et Février 2018 a permis de vérifier, via l'utilisation d'un étalon commun d'intercalibration (ICM), si la pratique d'évaluation proposée dans notre nouveau dispositif national s'alignait ou non avec celle des autres Pays-Membres participants, à commencer par le calage du niveau de référence par défaut adopté pour chaque type et en mettant un focus particulier sur la limite inférieure de Bon Etat.

Suite à l'acceptation du rapport final d'intercalibration (résultats considérés comme OK pour notre pays), le nouveau dispositif d'évaluation des TGCE de France, vérifié DCE-compatible, a déjà fait l'objet d'une publication au JORF (nouvel Arrêté Evaluation signé en date du 27-07-2018) pour une entrée en application à partir du 1er Janvier 2019, en phase avec le début du 2^{ème} demi-PGME 2019-2021.

5.3. Examen des données existantes et repérage des cas particuliers

Le bilan quantitatif des données disponibles actualisé en Juillet 2017, en termes de nombre de sites et en nombre de relevés plus ou moins utilisables en doublets de données (chimie + assemblages diatomiques), a été récapitulé dans le Tableau 1, ainsi que les caractéristiques du jeu de données français soumis à l'intercalibration (qui avait dû être fourni dès septembre 2015, en fonction du planning de l'exercice).

2 niveaux différents de requêtes ont été effectués :

- l'une portant sur tous les sites faisant l'objet de suivis diatomées et à BV intégré > 10 000 km² (en conformité avec le cadrage typologique Européen du GIG Large Rivers),
- l'autre portant sur tous les sites > 8 000 km². Cette option pourrait être porteuse de sens, dans nos conditions nationales, dans l'éventualité de réviser l'évaluation des plus grands affluents des grands fleuves Français dans une cohérence d'évaluation entre amont et aval de leur confluence avec le drain principal. En effet, ce sont déjà de grands hydrosystèmes sans aucune situation relictuelle de référence vraie valide.

La Figure 32 représente, sur une base géographique, l'emplacement des sites TGCE disponibles sur le réseau fluvial Français (linéaires représenté à partir d'une surface intégrée minimale de 8 000 km² de BV intégré).

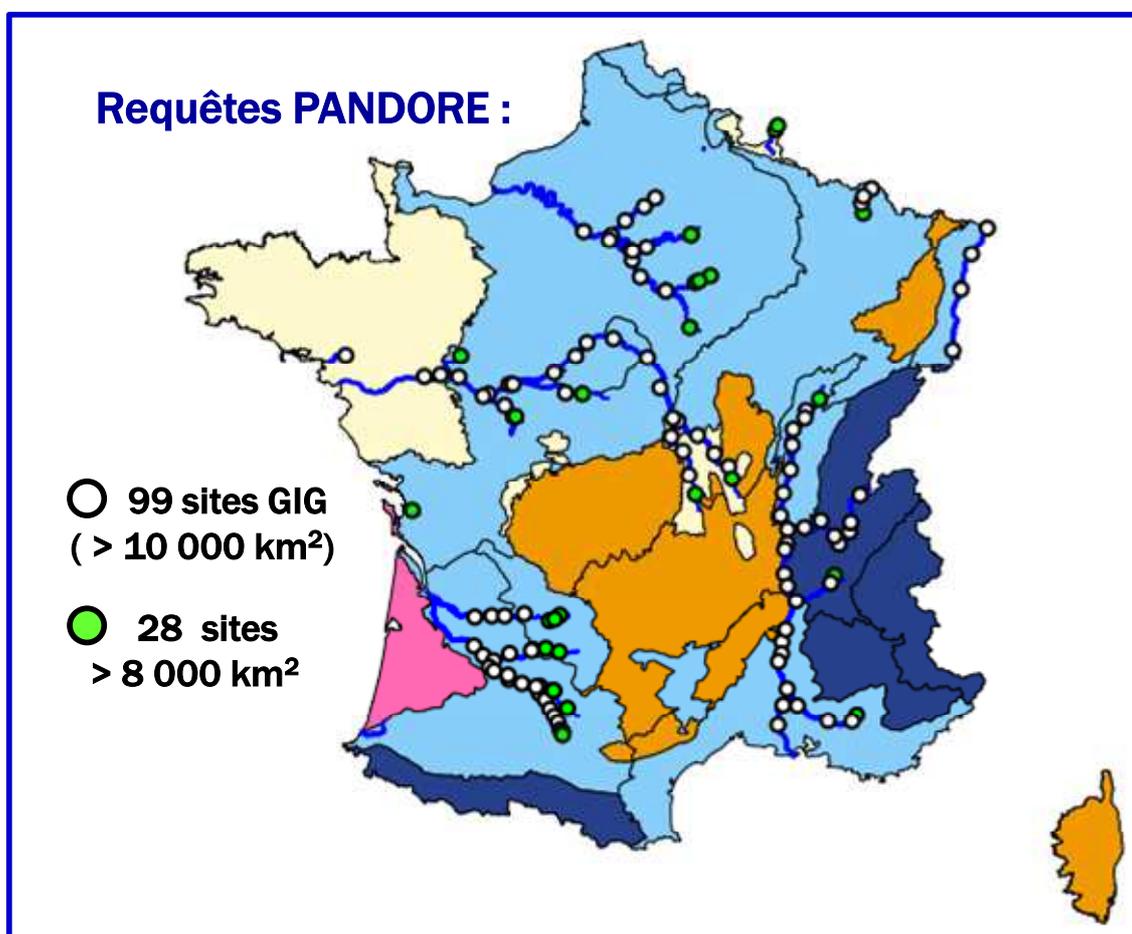


Figure 32 : Bilan des 127 sites nationaux TGCE associant descripteurs abiotiques chimiques et relevés diatomiques

99 sites de réseaux (figurés en blanc) présentent une surface intégrée conforme à la typologie européenne du GIG « Large Rivers » (> 10 000 km²),

28 sites supplémentaires (figurés en vert) présentent une surface intégrée un peu plus faible (> 8 000 km²), mais pourraient correspondre à un scénario d'évaluation intéressant dans notre contexte national.

En effet, cette taille intégrée de bassin versant modérément réduite intègre quelques gros affluents avant leur confluence avec le drain principal, ainsi que la Charente (fleuve côtier) et la Meuse (cours d'eau transfrontalier) et pourrait aussi concerner l'Adour un peu avant sa zone d'influence haline si une station aval intégrait des suivis diatomiques (ce qui n'est pas le cas actuellement). Au total, le gisement de données disponibles concerne 127 sites nationaux présentant, selon le cas, entre 1 et 6 relevés diatomiques avec couplage chimie sur la période.

Vu l'aspect plutôt insuffisant des référentiels TGCE disponibles pour la proposition et le calage des dispositifs antérieurs d'évaluation, le contenu des Arrêtés d'Evaluation publiés jusqu'à ce jour «faute de mieux» pouvait, dans le cas de ces très grands cours d'eau, présenter certains problèmes ou aspects non-satisfaisants.

A ce stade, dans le cadre de la présente étude nationale, il était donc utile que la situation géo-climatique de ces sites et leur implantation précise dans un réseau hydrographique donné fassent l'objet d'un ré-examen plus attentif afin de repérer et d'expertiser les particularités pouvant conduire à des anomalies d'évaluation. Il convient notamment de ré-examiner la pertinence du rattachement typologique aux régions naturelles diatomées, qui avait à l'époque été mis en place selon des critères d'application relativement automatiques ; ou bien, dans d'autres cas, l'adéquation éventuellement défectueuse du niveau de grille d'évaluation utilisée en fonction d'une configuration longitudinale particulière du drain principal et de ses confluences avec des affluents principaux.

Enfin, la dernière réactualisation de cet Arrêté Evaluation (27/07/2015) présentait encore une lacune complète d'évaluation des très grands cours d'eau d'origine Alpine (le Rhône, le Rhin), qui doit désormais être comblée de façon judicieuse.

La **Figure 33** ci-dessous différencie les sites (cercles à bordure noire) pour lesquels l'affectation automatique à la région diatomées traversée ne pose a priori pas de problème et ceux ou, pour différentes raisons, l'affectation antérieure pose question et semble pour le moins à ré-examiner, puis à modifier le cas échéant.

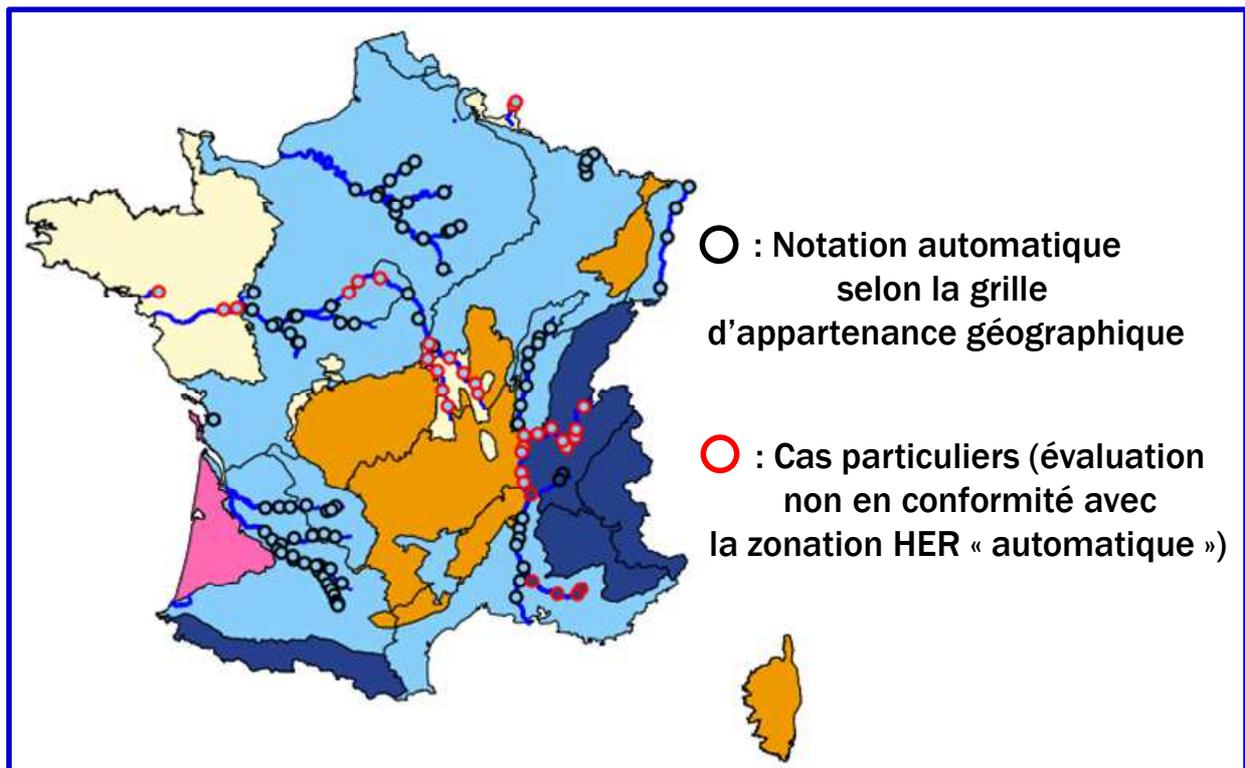


Figure 33 : Classification automatique de la grille d'évaluation selon la Région Diatomées d'appartenance ; illustration des cas spécifiques à expertiser (en rouge)

Les principaux sites posant question sur la stratégie nationale d'évaluation sont les suivants :

- a) Le **tronçon du Rhône automatiquement rattaché à la région Alpine** (Région Diatomique 5 de la Figure 29), de son entrée sur le territoire Français dans le secteur de Pougny jusqu'à son passage en zone méditerranéenne de plaine (Région Diatomique 4).

Deux problèmes se posent sur ce tronçon : **1)** L'évaluation du parcours français du Rhône sur son tronçon amont, après l'exutoire du lac Léman et jusqu'à l'agglomération de Lyon, n'était pas encore réglé dans le dernier Arrêté Evaluation (classification en TGCA, très Grand Cours d'Eau Alpin, mais case grisée avec dièse). Il est en tout cas certain que la grille de Région Diatomées n°5 (Figure 27) est trop sévère pour un cours d'eau de cette taille ayant de plus traversé de bout en bout le Lac Léman, qui est l'un des plus grands lacs d'Europe. **2)** L'autre problème concernant ce même tronçon du Rhône, mais cette fois à partir de l'agglomération de Lyon et avant son passage dans la Région Diatomées n°4 (passage dans le domaine Méditerranéen, contexte de plaine plutôt minéralisé), réside dans un changement notable de ses caractéristiques hydrochimiques à partir de sa confluence avec la Saône à Lyon, importante rivière de plaine très minéralisée.

b) La **Durance**, cours d'eau sous forte influence alpine, qui est actuellement évaluée dans le cadre de la Région Diatomées n°4 « Plaines, forte minéralisation », alors que l'**Isère** est évaluée avec la grille de la Région Diatomées n°5 « Zone Alpine », beaucoup plus sévère. Or cette différence de système de notation entre 2 rivières d'origine comparable n'est pas justifiée. De plus, la station la plus aval de l'Isère passe actuellement sur une grille d'évaluation plus laxiste pour une part modeste de son cours (passage en zone de plaine minéralisée du domaine méditerranéen). Or l'instauration de cette discontinuité d'évaluation est difficile à justifier dans le contexte où le passage dans cette nouvelle région naturelle, quelques kilomètres seulement avant sa confluence avec le Rhône, n'a pas le temps de modifier les caractéristiques hydrochimiques et thermiques de la rivière, et où cette station est de plus la mieux notée de tout le parcours TGCE de l'Isère.

c) Problème de l'**Allier** et de la **Loire** avec le passage automatique de la Région Diatomées n°3 à l'amont, qui procure une évaluation assez sévère (mais à ce stade, ce ne sont pas encore des TGCE...) à la Région Diatomées n°2 pendant le passage dans les dépressions sédimentaires de Nord-Massif Central ou de Sologne (grille du niveau de sévérité le plus bas), puis à nouveau, complètement à l'aval, lors de l'arrivée dans le secteur Sud-Bretagne. Dans le secteur intermédiaire, c'est une grille de notation n°4 (cours d'eau de plaine, eaux minéralisées) qui s'appliquerait, moins sévère que la grille 3 mais plus sévère que la grille 2. Cette stratégie d'évaluation génère une incohérence sur le plan du continuum longitudinal d'aussi grands cours d'eau (forcément traversants par rapport à une zone naturelle donnée...). Dans un tel cas, il est préférable d'adopter une solution permettant une évaluation logique, lissée par rapport à l'hydrochimie nécessairement composite au niveau de si larges bassins versants et par rapport à la place des sites dans le continuum amont-aval, et sans discontinuité artificielle de notation qui serait uniquement générée par un changement de zone naturelle et de grille d'évaluation, alors que la courte traversée d'une autre HER n'aura pas encore eu le temps d'influer significativement sur l'hydrochimie de fond du TGCE.

d) Grille de notation différente et à différentiel d'évaluation discutable entre la **Meuse** et la **Moselle**.

e) Cas de la **Vilaine aval** à ré-expertiser par sécurité.

Hormis ces cas qui précèdent, pour lesquels le problème réside dans une adéquation discutable entre localisation du site et grilles d'évaluation correspondantes par affectation géographique automatique, il va aussi falloir proposer un dispositif d'évaluation pour le **Rhin**, cours d'eau rattaché à un type national (TTGA) actuellement non-évalué, qui prend sa source et s'écoule en Suisse, puis qui devient fleuve transfrontalier entre ce pays et le Liechtenstein puis l'Allemagne, pays où il traverse aussi un grand lac (le Lac de Constance), avant de débiter son parcours frontalier entre France et Allemagne au long de la Plaine d'Alsace. Le cas de ce fleuve est *a priori* assez comparable à celui du Rhône dans son secteur en amont de Lyon, de l'aval de sa traversée du Léman jusqu'à sa confluence avec la Saône. Sur le plan réglementaire, ces deux cours d'eau se trouvent actuellement dans le même cas (classification en TGCA, Très Grands Fleuves Alpains, case de couleur grisée contenant un dièse, signalant un contexte d'évaluation encore non défini). Il va donc falloir étudier plus avant les données et proposer un dispositif d'évaluation si possible cohérent entre le Rhône sur le secteur amont de Lyon et le Rhin dans la partie plaine d'Alsace.

Le cas du Rhône a été traité de la façon suivante. Actuellement, selon la typologie nationale, ce fleuve se rattache géographiquement à 2 régions diatomées (la Région 5 ou région Alpine, la Région 4 de plaine minéralisée, en zone Méditerranéenne, cf. Figure 29). Or 2 problèmes sont à résoudre :

- Il convient d'établir un niveau de notation correct pour le tronçon qui court depuis l'exutoire du Lac Léman jusqu'à Lyon. Si le rattachement au secteur alpin est logique géographiquement et hydrologiquement, la grille 5 diatomées actuellement existante pour des cours d'eau alpins plus petits ne serait pas judicieuse sur le Rhône dans ce secteur car elle serait exagérément sévère, compte tenu de la dimension déjà importante du BV cumulé de cet hydrosystème et du fait qu'il a traversé un grand lac alpin au volume très important (le Léman) avant de reprendre son cours fluvial. Jusqu'à présent, la case correspondante du dernier Arrêté était en grisé et avec dièse (pas de grille de notation proposée), et la grille alpine actuelle pour de plus petits hydrosystèmes (grille de la Région 5) serait manifestement trop sévère un cours d'eau de ce type.

- Un 2^{ème} problème vis-à-vis de l'évaluation de ce tronçon sub-alpin se situe depuis l'aval de la confluence avec la Saône, grosse rivière très minéralisée et plutôt de plaine, qui change très significativement l'hydrochimie du Rhône à partir de Lyon, jusqu'à la confluence avec l'Isère. Sur ce tronçon, le Rhône était aussi classifié en TGCA et n'était pas évaluable jusqu'à présent. Il va falloir proposer une pratique d'évaluation adaptée tenant bien compte du contexte longitudinal de ce fleuve et de l'influence de sa confluence avec la Saône, actuellement évaluée pour sa part avec la grille de la Région 4 « Plaines minéralisées ». La logique conduirait à ce qu'il en devienne de même pour le Rhône à l'aval de cette confluence.

Ensuite, par référence à la typo nationale et à partir de l'aval de sa confluence avec l'Isère, le Rhône était effectivement évalué selon une grille de rivière de plaine minéralisée en zone méditerranéenne (Grille 4). Cette construction semblait assez logique sur le plan géochimique et bioclimatique mais pose cependant un problème de représentativité d'évaluation vis-à-vis des situations de référence trouvées dans cette Région Diatomées (qui appartiennent à des hydrosystèmes beaucoup plus petit que le Rhône aval).

5.4. Premier prototype de dispositif Français TGCE proposé

5.4.1. Liaison avec les échéanciers de travail du GIG Large Rivers

Durant le 2^{ème} semestre 2015, le démarrage de l'exercice européen d'intercalibration TGCE dans les meilleurs délais possibles a occasionné un niveau de contrainte important vis-à-vis des échéanciers. En effet, il fallait proposer le plus rapidement possible un prototype de système national d'évaluation présentant une bonne cohérence interne et ayant les meilleures chances d'être maintenues par la suite (sauf écart injustifié par rapport aux autres Pays-Membres, qui pourrait nécessiter ensuite un éventuel réalignement).

Pour que cet exercice puisse débiter, la coordinatrice de l'exercice, Dasa HLUBIKOVA, a demandé aux Etats-Membres voulant participer, sous délai contraint :

- de décrire avec précision leur système national à intercaler et ses principes de fonctionnement. Nous avons donc renvoyé au GIG, en date du 15-12-2015, une note nationale d'expertise rédigée en anglais qui décrivait le système d'évaluation français et son mode de fonctionnement (cf. **REF. 4**) ;
- d'envoyer les données d'évaluation diatomique du jeu de données national soumis à cette intercalibration. Cet envoi devait inclure le calcul d'EQR par relevé diatomique dans la bonne zone géographique et selon la bonne grille d'évaluation, ce qui présupposait d'avoir déjà défini un système d'évaluation-candidat et d'avoir travaillé à la résolution des quelques problèmes diagnostiqués et listés au Chapitre 5-3 qui précède.

Pour pouvoir nous acquitter de ces requêtes dans les temps, il a fallu réaliser dans l'urgence un travail permettant d'effectuer une proposition de dispositif national révisé spécifiquement dédiée TGCE, et envoyer aussi au GIG le jeu de données français TGCE « définitif », incluant les modifications de calculs d'EQRs (requis pour le 15 à 20 Décembre 2015).

En fonction de cette première base et selon résultats nationaux obtenus, plusieurs itérations ont ensuite été nécessaires pour affiner la proposition et entrer dans les biais jugés tolérables pour l'exercice.

5.4.2. Réadaptation du dispositif national au cas des TGCE : 1^{ère} proposition

La première adaptation de grilles d'évaluation proposée pour évaluer plus judicieusement les TGCE de France (1^{er} prototype de dispositif national, soumis au 12 décembre 2015) est présentée en **Figure 34** page suivante.

Sur notre territoire de France métropolitaine comme pour beaucoup d'autres Pays-Membres européens, plus aucun site de référence vraie valide n'existe sur très grands cours d'eau. Beaucoup des sites de référence repérés au niveau national se trouvent sur des hydrosystèmes petits à moyens, souvent adossés à des reliefs et plutôt situés en périphérie des grands bassins hydrographiques. Et même dans de telles situations positionnées très en amont de beaucoup de bassins versants, assez souvent, des pressions anthropiques notables peuvent déjà s'exercer (implantation de villages et / ou élevages de montagne, stations de ski, etc...). On ne va donc trouver des situations de référence vraie valides que dans d'assez rares secteurs situés dans des environnements particulièrement préservés.

a) Grilles d'évaluation en notes d'IBD 2007

IBD ₂₀₀₇	REF	H/G	G/M	M/P	P/B	MIN
Type 1	18,4	17,538	15,452	12,878	8,828	5,0
Type 2 (*)	17,4	16,416	13,792	10,020	5,920	1,0
Type 3	18,0	18,180	15,820	12,780	8,280	5,0
Type 4 (*)	18,1	17,074	14,338	10,405	6,130	1,0
Type 5 (*)	20,0	19,100	16,700	13,250	9,500	5,0
Type 6 (*)	18,6	17,544	14,728	10,680	6,280	1,0

(*) : Grilles utilisées sur TGCE (niveaux de référence par défaut)

b) Grille unique d'évaluation en EQRs

Biological Quality Element	Biological Index	Reference (median)	Ecological status class limits (in IBD-EQRs)				
			H/G	G/M	M/P	P/B	MIN
Phytobenthos (diatoms)	IBD ₂₀₀₇	1	0,94	0,78	0,55	0,30	0

Figure 34 : 1^{ère} proposition de grilles d'évaluation de l'Etat Ecologique pour les Très Grands Cours d'Eau : a) en notes d'IBD 2007 (arrondies au 1/1000^{ème}) ; b) en grille unique d'EQRs

Afin d'acquiescer le statut de références, les sites-candidats repérés par les Agences de Bassins et/ou les Services de l'Environnement vont devoir être soumis à diverses vérifications de conformité édictées par l'U.E. via la Guidance REFCOND (REF. 11 : Refcond Guidance, 2003), qui permettent de valider ce statut de référence vraie (ensemble des critères respectés à la fois vis-à-vis de l'occupation du sol et sur les analyses chimiques de l'eau) ou de leur conférer le statut de référence par défaut (respect de la plupart des critères imposés et mise en défaut, au moins à certains moments, d'autres critères).

Cependant, à partir du moment où on s'intéresse à des bassins versants intégrés de grande taille de type TGCE (ici, surface cumulée de BV > 8 000 à 10 000 km²), avec la densification de population et d'activités humaines dont sont porteurs les territoires, diverses catégories de pressions anthropiques deviennent inévitables et provoquent forcément des effets non négligeables à un moment ou à un autre. Dans ces conditions, il n'est plus possible de respecter le même principe de construction de grilles d'évaluation que pour de plus petits cours d'eau, qui avaient été calculées à partir du niveau de référence médian réellement observé pour une région diatomée donnée. Pour cette première proposition, nous avons essayé de faire preuve de pragmatisme, en nous appuyant le plus souvent possible sur le système d'évaluation déjà proposé pour les autres types de cours d'eau (utilisation chaque fois que possible de grilles pré-existantes, utilisation des mêmes limites de classes de qualité en EQR), et en ne complexifiant le dispositif qu'à la marge pour permettre d'évaluer sur une base équitable les quelques TGCE qui ne pouvaient l'être sur la base d'aucune des grilles déjà utilisées.

Dans tous les cas, les niveaux de référence de ces grilles ne résultent pas d'une statistique de relevés observés sur de vrais sites de TGCE de référence (alors que, pour des cours d'eau de taille plus modeste, il s'agissait de la médiane des vraies notes de référence), mais à des niveaux de référence projetés sur des cours d'eau déjà sous le coup d'altérations et de baisses de notes, ou *a minima* à des conditions de moindre effet de l'anthropisation (= *least disturbed condition*). Il faut préciser qu'en France, ce dernier cas n'est plus rencontré que sur de rares rivières de cette échelle de taille (Durance, Dordogne...) suffisamment préservées pour que les notations indicelles puissent encore atteindre les notes maximales de façon assez courante.

Par rapport aux types naturels de flore diatomique présents sur des systèmes plus petits, il faut noter que le seuil minimal de taille requis ici pour obtenir le jeu de données TGCE (> 8 000 km²) implique déjà une baisse sensible d'altitude des sites concernés et s'accompagne très souvent d'une composition hydrochimique naturelle nettement plus composite que pour de plus petits hydrosystèmes. Ainsi il n'est plus possible, à cette échelle de taille, de rencontrer des conditions hydrochimiques naturelles très particulières et circonscrites, par exemple celles correspondant au massif sableux des Landes de Gascogne (eaux très peu minéralisées et acides, flores naturelles acidobiontes).

Aussi, pour les cours d'eau correspondant au seuil de taille minimal requis par le GIG (>10 000 km²), il devient donc tout-à-fait exceptionnel, au niveau de tout notre territoire Français, de les voir se classer dans la typologie de faible alcalinité (<0,5 meq Ca CO₃) alors que, même si cette catégorie de cours d'eau est relativement minoritaire en France du fait de la large couverture par des affleurements ou bassins sédimentaires calcaires, ce cas de figure de « cours d'eau peu minéralisés » n'est pas rare pour des hydrosystèmes plus petits. Dans cette catégorie de taille, seule la Vilaine aval rentre dans ce seuil et d'autres rivières modestement minéralisées (ex : Allier) le dépassent déjà légèrement sur la base de la valeur moyenne annuelle.

Enfin, il est manifeste que la plupart des TGCE ne pourraient pas être évalués de façon judicieuse avec des grilles aussi sévères que pour des cours d'eau plus petits, sauf cas particulier de certains affluents arrivant directement de zones de hautes montagnes.

Compte-tenu de ce qui précède, par rapport aux grilles pré-existantes dans le système national d'évaluation diatomique des cours d'eau, 2 grilles deviennent sans objet pour des systèmes aussi grands et composites (cf Figure 34, partie a) :

- la grille de Région Diatomique 1 (flore acidobiontes des cours d'eau acides des landes)
- la grille de Région 3 (flore naturelles des hautes-terres du Massif Central, sur eaux faiblement minéralisées, dont le niveau de référence est trop élevé pour presque tous les TGCE).

Les grilles suivantes ont été utilisées pour évaluer les TGCE dans le cadre de la 1^{ère} proposition de dispositif :

- La grille de Région 2 ne peut être utilisée que de façon anecdotique (cas unique de la Vilaine aval, issue de la couverture hercynienne du socle Armoricaïn), compte-tenu des niveaux composites d'alcalinité atteints par la plupart des cours d'eau de cet ordre de taille,.
- La grille de Région 4 (plaines minéralisées) est convenable pour la plupart des TGCE s'écoulant dans un contexte carbonaté de moyenne à faible altitude,
- La grille de Région 5 (reliefs alpins), qui est basée sur des niveaux de référence élevés et sur l'impossibilité d'atteindre les plus basses notes d'IBD 2007 pour cause d'hydrodynamique et de dilution élevées, pourrait *a priori* être propagée vers l'aval pour 2 affluents alpins du Rhône : l'Isère et la Durance, directement adossées à un réseau hydrographique montagnard.

Il a été jugé nécessaire de rajouter à ce dispositif, basé sur des grilles d'évaluation pré-existantes, une nouvelle grille N° 6, moins sévère, pour permettre d'évaluer sur une base judicieuse de grands hydrosystèmes issus de zones montagneuses, mais plus éloignés des reliefs et aux bassins versants plus intégrateurs. Les TGCE concernés se situent tous à l'aval de systèmes lénitiques à gros volume interne, ce qui nécessite aussi une adaptation à la baisse du niveau de référence par défaut adopté. Les 3 hydrosystèmes concernés par l'application de la grille 6 sont le Rhône depuis la sortie du Lac Léman jusqu'à Lyon, le Rhin, qui a traversé le Lac de Constance, et la Dordogne dont le système hydrographique amont, sur drain principal et sur ses affluents principaux, traverse de grosses retenues (Bort-les-Orgues, Entraygues-sur-Truyère etc...).

Jusqu'à présent, dans les constructions de dispositif national d'évaluation concernant le maillon diatomique, nous avons considéré que les particularités des régions naturelles (notamment niveau de référence et niveau le plus bas de notation indicielle qu'il est possible d'atteindre dans un ensemble naturel donné) devaient servir au calage régionalisé de niveau des grilles, mais que la grille d'évaluation en EQR devait être unique et donner une équivalence unique du seuil de Bon Etat par rapport au niveau de référence adéquat (tous les Etats-Membres ne sont pas forcément partis sur le même schéma...). Aussi, dans cette première proposition de dispositif d'évaluation TGCE, nous avons conservé le principe de fonctionnement basé sur une **grille unique d'évaluation en EQR** (partie b de la Figure 34).

En résumé, c'est sur la base du système de grilles tel que présenté dans cette Figure 34 qu'ont été effectués les calculs d'EQRs contenus dans le jeu de données français fourni en décembre 2015 pour l'exercice d'intercalibration 'Large Rivers ». Ce jeu de données contenait une sélection nationale de 221 relevés diatomiques prélevés sur 95 sites situés au long de TGCE de plus de 10 000 km².

5.4.3. Résolution des cas particuliers et 1^{er} dispositif d'évaluation proposé :

Les cas particuliers d'évaluation à ré-expertiser, suite à l'intervention pas toujours adéquate d'une affectation « automatique » de grille d'évaluation par rapport à l'HER de rattachement du site à évaluer, ont été illustrés sur la **Figure 33** présentée en p. 60 (sites cerclés de rouge).

Sur cette base, une expertise a été traitée bassin par bassin et les principaux éléments de décision entrant en jeu pour cette première proposition de dispositif sont illustrés dans les figures suivantes (de 35 à 38). A noter que les stations de réseaux, ainsi que leur classe d'évaluation présentée dans ces figures, ont été construites à partir de résultats d'évaluation issue de la Campagne 2012 des réseaux d'observation, dont le dispositif était à peu près complet et plutôt représentatif des réseaux actuels de surveillance en place sur TGCE.

a) Bassin du Rhône :

Le Rhône dans sa partie située entre l'aval du Lac Léman et sa confluence avec l'Isère (aval de Jons) était considéré dans la typologie nationale comme un TGCA (Très Grand Cours d'Eau Alpin), mais ne faisait pas encore l'objet d'une évaluation sur la base du compartiment diatomique (case en grisé avec dièse).

Les particularités de son cours sur ce secteur sont les suivantes :

- Son arrivée en France (secteur de Pougny) s'effectue en aval du lac Léman, sur une surface cumulée de bassin versant (incluant le parcours fluvial initial en Suisse et la surface de la cuvette en eau du lac Léman) de 10 370 km².
- Après son parcours dans le Jura français, il arrive dans l'agglomération de Lyon où se situe sa confluence avec la Saône. Cette confluence marque un changement typologique assez important (outre le changement hydrologique lié à l'apport d'une rivière de plaine, l'hydrochimie de base de l'eau subit un accroissement net d'alcalinité du à cette rivière très minéralisée avec des influences halines).
- A l'aval de cette confluence, ce fleuve s'écoule toujours dans une HER rattachée à l'ensemble alpin jusqu'à l'aval de Jons, où intervient sa confluence avec l'Isère.
- Cependant, sur cette partie à l'aval de Lyon, le Rhône ne peut plus être considéré comme typologiquement identique à la partie Jurassienne de son cours (modification de son hydrochimie suite à sa confluence avec la Saône, bassin versant intégré plus composite et d'ordre de taille sensiblement accru, aménagements anthropiques physiques de son cours...).
- -Dans le secteur de la confluence avec l'Isère, sur la base du zonage des HER1, il passe dans le domaine des cours d'eau de plaines minéralisées sous influence méditerranéenne.
- Il faut préciser aussi que l'affectation typologique automatique sur la base de l'HER traversée au site pose également problème pour la Saône, qui entre dans une HER alpine à son site le plus aval (Lyon) alors qu'elle est considérée comme rivière carbonatée de plaine sur tout son cours de TGCE plus à l'amont.
- Enfin, la différence de traitement actuelle pour l'Isère et la Durance, basée sur la projection géographique des HER1 alors qu'elles sont toutes deux issues de la haute montagne alpine, n'est pas justifiée.

Dans le dispositif actuel, l'Isère est évaluée comme un TGCE alpin, ce qui est logique, et passe en domaine méditerranéen sur sa station de réseau la plus aval (Châteauneuf-sur-Isère), juste en amont de sa confluence avec le Rhône. Mais ce court passage dans une HER de plaine calcaire méditerranéenne, à faible distance avant sa confluence, n'a le temps de changer ni son régime, ni son hydrochimie de base, ni ses caractéristiques thermiques *in situ*. Il vaut donc mieux l'évaluer sur la même grille jusqu'à sa confluence avec le Rhône pour éviter de générer une discontinuité longitudinale d'évaluation.

De son côté, la portion de la Durance classée en TGCE est géo-localisée automatiquement en zone méditerranéenne et induirait l'emploi d'une grille d'évaluation plus laxiste que celle de l'Isère (la grille 4 au lieu de la 5), ce qui ne se justifie ni en fonction du bassin versant d'origine de ces cours d'eau (zone de montagne alpine dans les 2 cas), ni en fonction du niveau de référence à considérer. En effet, l'Isère est sensiblement plus altérée que la Durance qui, malgré sa taille et les problèmes importants d'abstraction d'eau qu'elle subit, présente une qualité d'eau plutôt bonne et obtient des notes d'indice diatomique élevées, proches des niveaux de référence pour la zone alpine. Sur ce maillon en tout cas, l'emploi d'une grille d'évaluation de type plaine méditerranéenne ne se justifie pas et créerait une incohérence avec l'évaluation de l'Isère.

La proposition de dispositif d'évaluation pour cet hydrosystème du Rhône et affluents, récapitulée en **Figure 35** (page suivante), se présente donc ainsi :

- évaluer la Saône sur tout son cours TGCE, y compris sur son court parcours dans l'HER Alpes, sur la Grille 4 (rivières de plaine minéralisées),
- évaluer l'Isère jusqu'à sa confluence avec le Rhône et la Durance jusqu'à sa confluence avec le Rhône sur la base de la Grille 5 (rivières alpines),

- évaluer le Rhône sur son parcours Jurassien, jusqu'à sa confluence avec la Saône à Lyon, sur la base de la nouvelle Grille 6,
- évaluer le Rhône à partir de sa confluence avec la Saône à Lyon jusqu'à son embouchure sur la base de la même Grille 4 (d'abord cours d'eau de type « rivière de plaine minéralisée » depuis l'aval de sa confluence avec la Saône jusqu'à la confluence avec l'Isère, puis cours d'eau rattaché à l'HER « plaines calcaires méditerranéennes » de la confluence avec l'Isère jusqu'à son embouchure sur la Méditerranée).

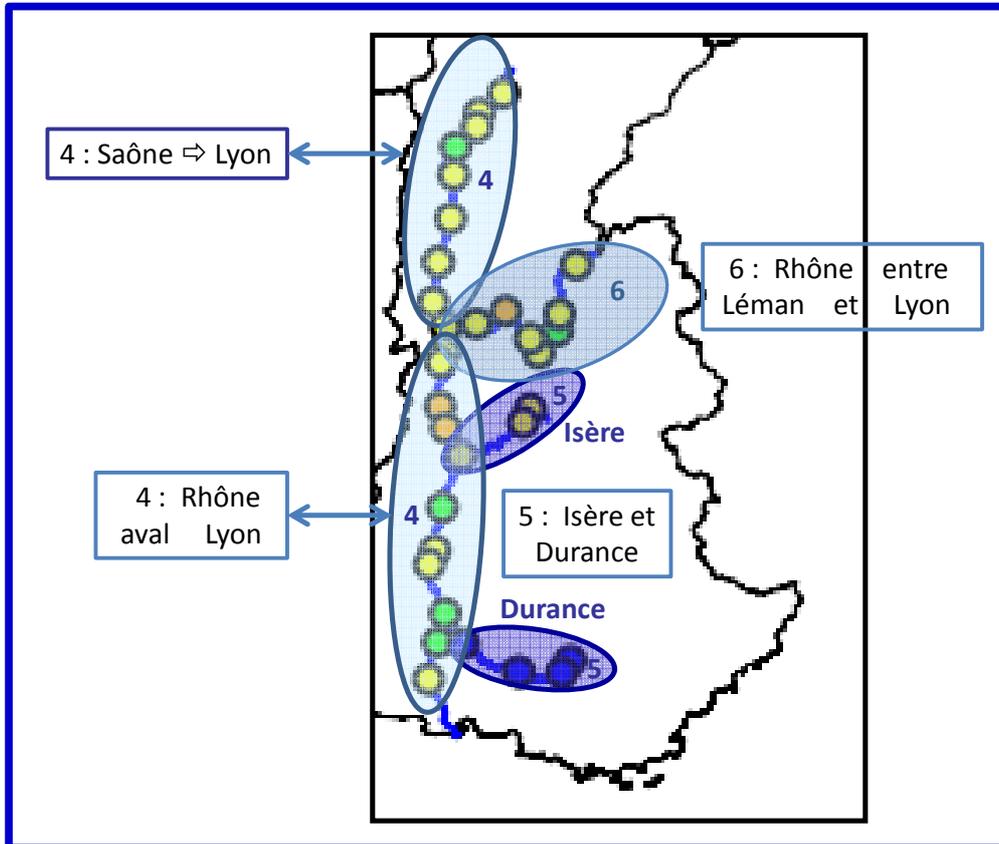


Figure 35 : 1^{er} dispositif d'évaluation proposé pour les TGCE du bassin hydrographique du Rhône

b) TGCE du Nord-Est de la France :

Les 3 hydrosystèmes classés TGCE dans ce secteur sont le Rhin, la Meuse et la Moselle (cf. **Figure 36** en page suivante).

Bien que présentant déjà une surface cumulée sensiblement plus importante que le Rhône dans son secteur Jurassien, le Rhin, qui cumule déjà 37 000 km² à son entrée en France dans la Plaine d'Alsace (http://www.eau-rhin-meuse.fr/connaissance_du_bassin_rhin_meuse), présente d'assez grandes similitudes avec le Rhône à l'amont de Lyon.

En effet, le cours amont du Rhin est implanté en Suisse, où il présente des secteurs frontaliers avec le Liechtenstein et l'Allemagne. De plus, ce fleuve a aussi dû traverser un grand lac alpin (le lac de Constance) en amont de son arrivée en plaine d'Alsace. Enfin, comme le Rhône, il a été classé dans le type TTGA mais comme lui, au niveau réglementaire, il n'était pas encore évaluable jusqu'à présent dans le dispositif national d'évaluation (case grisée avec un dièse).

Pour pouvoir envoyer le jeu de données national avec les calculs d'EQRs, la proposition a été d'évaluer le Rhin sur la base de la nouvelle grille proposée pour les grands TGCE originaires de zones montagneuses (Grille 6).

Sur la base de l'attribution automatique de typologie en fonction de l'HER d'appartenance, le cas de la Meuse et de la Moselle, rivières de Lorraine assez proches géographiquement, pose question.

En effet, la Moselle est une rivière très minéralisée (Cond. El. de 1 500 à 1 800 $\mu\text{S}/\text{cm}$) subissant, avec certains de ses affluents (la Seille...), des influences halines à la fois naturelles (origine géologique) et anthropiques (industries chimiques). Cette rivière est logiquement évaluée sur la base de la Grille 4 Diatomées (cours d'eau minéralisés de plaine).

Par contre, la Meuse, rivière voisine ne subissant pas d'influence haline, est aussi fortement minéralisée et carbonatée (Cond. El. de l'ordre de 500 $\mu\text{S}/\text{cm}$). Sur la base de son classement typologique automatique (passage dans les Ardennes), elle était évaluée de façon erronée comme TGCE sur la base de la Grille 2, grille peu sévère correspondant plutôt à des cours d'eau peu minéralisés sous influence géologique hercynienne, alors que presque tout son bassin versant amont est sous influence géologique de sédiments calcaires du Bassin Parisien. Comme, de plus, elle ne présente pas de problème de niveau d'évaluation (elle est plutôt de bonne qualité), et du fait de son statut de fleuve trans-frontalier, il vaut mieux adopter une stratégie d'évaluation conservatrice pour le pays situé à l'aval.

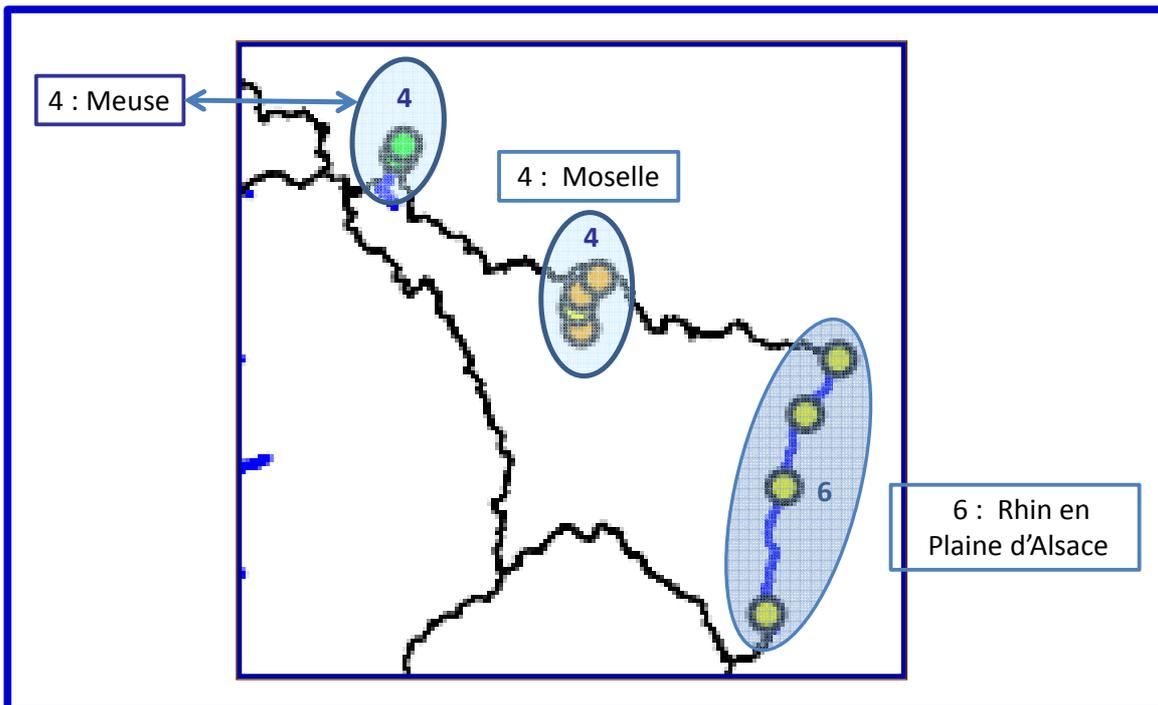


Figure 36 : 1^{er} dispositif d'évaluation proposé pour les TGCE du Nord-Est de la France

La proposition est donc de continuer d'évaluer la Meuse dans son secteur TGCE sur la base de la même grille de qualité que la Moselle et que son réseau amont + affluents, à savoir la Grille 4. Les calculs d'évaluation préalables à l'envoi du jeu de données à intercalibrer ont donc été faits dans ce sens.

c) Cours d'eau de l'Ouest de la France :

Les 3 réseaux hydrographiques TGCE de ce secteur sont la Seine, la Vilaine et la Loire (cf. **Figure 37** en page suivante).

La Seine et son réseau d'affluents est complètement originaire du contexte géologique sédimentaire du Bassin Parisien. Aussi, l'évaluation antérieurement pratiquée, qui correspondait à une affectation typologique d'évaluation sur la base de la Grille 4 Diatomées, ne pose pas question et sera maintenue.

Le cas de la Vilaine, assez atypique au niveau de sa classification, a fait l'objet d'un ré-examen. Tout le réseau hydrographique amont de cette rivière est originaire de formations peu minéralisées d'origine hercynienne (massif Armoricaïn) ou de ses contreforts à substrats géologiques également peu minéralisés (faible présence de calcium et de carbonates). Cette rivière est donc le seul TGCE de plus de 10 000 km^2 de France qui a été rattachée au type GIG correspondant aux rivières de faible alcalinité ($<0,5 \text{ meq Ca CO}_3 / \text{l}$). Cela justifie son évaluation selon la Grille 2 Diatomées (grille la moins sévère, réservée aux cours d'eau peu minéralisés de plaine). C'est le seul cours d'eau de cet ordre de taille évalué selon cette grille, qui concerne aussi des cours d'eau de plus faible taille originaires de substrats hercyniens (dont Massif Armoricaïn, dépressions sédimentaires Nord-Massif Central...).

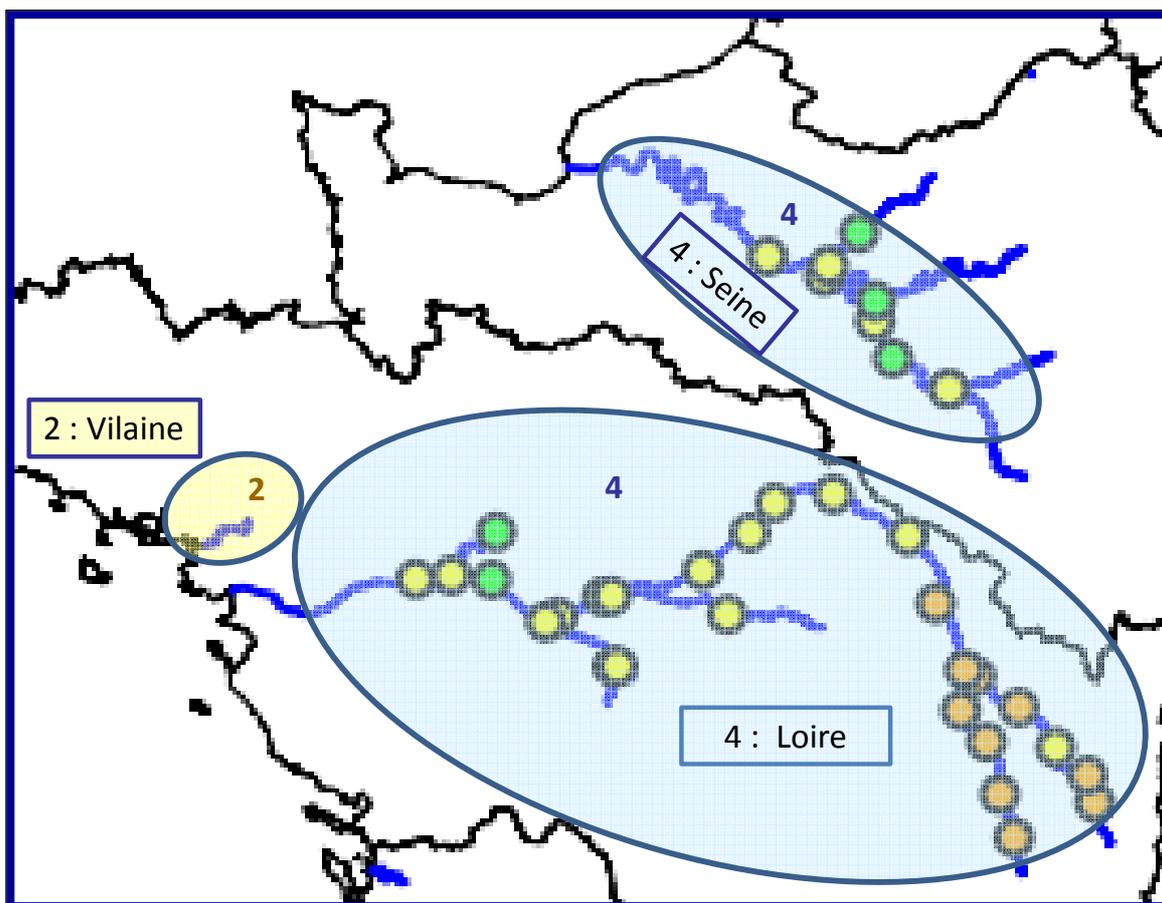


Figure 37 : 1^{er} dispositif d'évaluation proposé pour les TGCE de l'Ouest de la France

La Loire et son principal affluent l'Allier sont concernés par des problèmes d'affectation automatique de grille d'évaluation en fonction de l'HER traversée, qui posent des problèmes de zonation longitudinale et de cohérence d'évaluation amont-aval.

Ces 2 cours d'eau sont originaires des hautes terres du Massif Central et leur réseau hydrographique amont est évalué sur la base de la Grille 3, spécifique à cette région peu minéralisée d'altitude (cf Figure 30) et relativement sévère, mais non proposée à l'application pour les TGCE (cf Figure 34).

Leur arrivée dans la catégorie de taille TGCE coïncide avec leur arrivée dans les dépressions sédimentaires du Nord-Massif Central, qu'elles traversent assez rapidement et où, sur base d'affectation automatique, ils se verraient attribuer la Grille 2 diatomées, assez laxiste. Ensuite, ces cours d'eau arrivent en zone de plaine calcaire dans un système de grille notant à nouveau un peu plus sévèrement. Enfin, la Loire aval arrive en zone peu minéralisée de l'ensemble géologique Armoricaïn dans lequel s'appliquerait une nouvelle fois la Grille 2 diatomées, la moins sévère.

Il vaut mieux adopter des dispositions évitant, au long du même continuum Loire + Allier, des effets de discontinuités longitudinales d'évaluation peu justifiables (court passage, après une grille 3 sévère sur les hautes terres du Massif Central, à une grille 2 peu sévère, puis passage à une grille 3 à nouveau un peu plus sévère correspondant aux plaines carbonatées, pour terminer une nouvelle fois sur la grille 2 la moins sévère (rivières peu minéralisées de plaine) alors que la Loire moyenne et aval est sensiblement minéralisée.

La proposition de dispositif pour l'hydrosystème Loire telle que présentée en Figure 37, est donc de démarrer l'évaluation sur la base de la Grille Diatomées 4 (couleur bleue) dès l'arrivée dans les dépressions sédimentaires du Nord du Massif Central (grille de sévérité intermédiaire entre la Grille 3 et la Grille 2, dont le laxisme ne se justifierait pas vu le court parcours dans cette zone de ces longs cours d'eau traversants), de la conserver dans toute la zone carbonatée du Val de Loire (ce qui correspond à l'affectation typologique automatique), et de la prolonger jusqu'à la zone estuarienne (sur le plan hydrochimique, la Loire ne subit pratiquement pas l'influence de la zone Sud-Armoricaine qu'elle traverse, et reste une rivière assez fortement minéralisée jusqu'à son exutoire).

Ainsi, le système d'évaluation à mettre en œuvre pour la Loire et l'Allier est très simple pour les intervenants gestionnaires (2 grilles seulement à appliquer, une grille relativement sévère sur le secteur montagneux amont et une grille de plaines carbonatées depuis la sortie du Massif Central jusqu'à l'embouchure de la Loire) et ne génère pas d'effets artificiels de discontinuité longitudinale d'évaluation non basés sur les caractéristiques hydrochimiques réelles de cet hydrosystème.

d) Hydrosystème Garonne-Dordogne :

Les cours d'eau de cet hydrosystème qui atteignent le seuil minimal de taille les faisant entrer dans la catégorie TGCE (ici, au moins 8 000 km²) sont la Garonne, la Dordogne, le Tarn et le Lot (**Figure 38**).

A noter que, dans ce quart Sud-Ouest de la France, l'Adour sur sa partie aval entrerait aussi dans cette catégorie (portion figurée en bleu à l'aval de ce cours d'eau). Cependant, nos requêtes ne nous ont pas permis de trouver une station suffisamment aval couplant l'atteinte de ce seuil de taille et l'existence de relevés diatomiques sur la période.

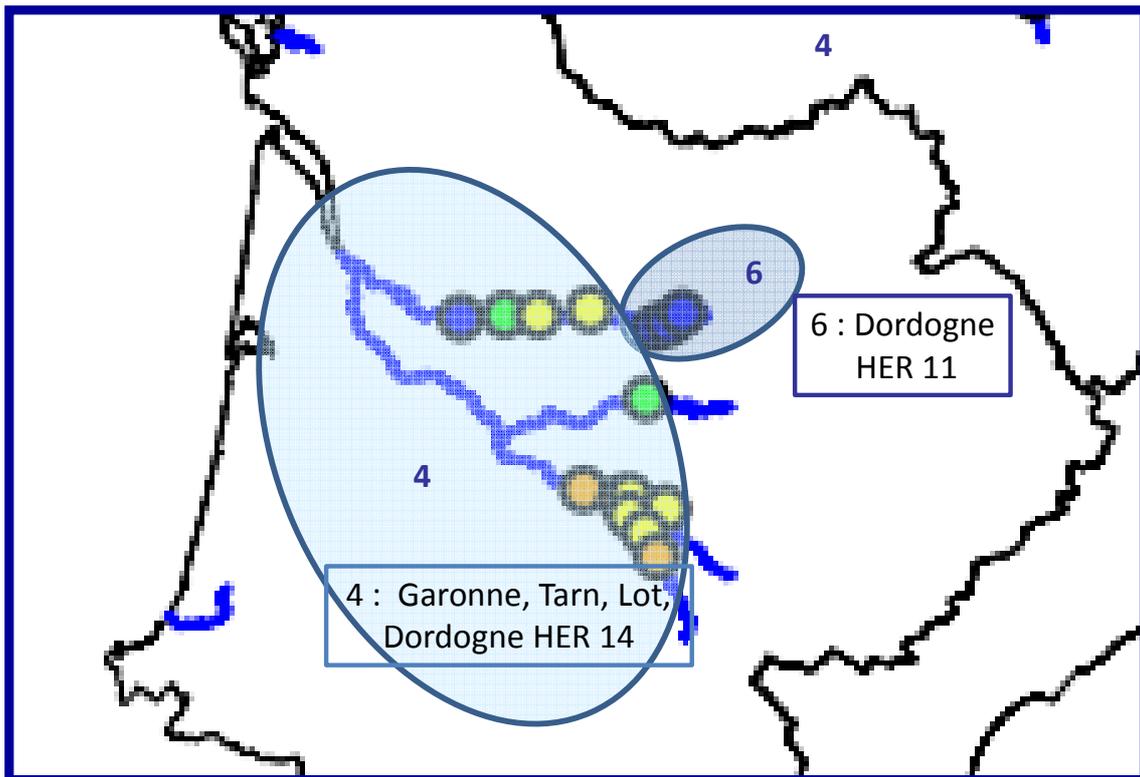


Figure 38 : 1^{er} dispositif d'évaluation proposé pour l'hydrosystème Garonne-Dordogne

Pour la Garonne, le Tarn, le Lot et la Dordogne aval à partir de son passage dans l'HER 14 « Coteaux Aquitains », l'évaluation sur base d'affectation typologique automatique ne pose pas de problème et ces TGCE peuvent être évalués de façon tout-à-fait valide sur la base de la Grille 4 Diatomées.

La Dordogne amont, avant d'atteindre la taille de TGCE, est évaluée à l'aide de la Grille 3 Diatomées, grille assez sévère correspondant aux cours d'eau de taille petite à grande aux eaux peu minéralisées provenant des hautes-terres du Massif Central. Ce principe pré-existant ne pose pas de problème pour ce cours d'eau dans ce secteur.

Ensuite, avant son arrivée dans les Coteaux Aquitains, la Dordogne dans son tronçon de l'HER 11 « Causses Aquitains » était pour l'instant évaluée selon la Grille 4 Diatomées. S'il n'est pas possible de continuer à évaluer cette rivière dans cette HER sur la base de la sévère Grille 3, dédiée à de plus petits hydrosystèmes du Massif Central et du fait de sa situation à l'aval de grosses retenues hydroélectriques (dont Bort-les-Orgues), il paraît par contre trop laxiste d'évaluer la Dordogne dans cette HER11 sur la base de cette Grille 4 « rivières minéralisées de plaine ». En effet, ce parcours s'écoule dans un secteur peu densifié et relativement bien préservée des impacts anthropiques touchant à la qualité des eaux, comme en témoignent les relevés indiciaires diatomiques qui peuvent atteindre de façon assez commune les plus hautes notes possibles d'IBD 2007 (proches de 20).

Une proposition *a priori* plus progressive et plus judicieuse pour évaluer la Dordogne sur son cours moyen en HER11 serait de s'appuyer sur la nouvelle Grille 6 Diatomées, également proposée à l'utilisation pour 2 TGCA (Très Grands Cours d'Eau Alpines), à savoir le Rhône sur son parcours dans le Jura français et le Rhin en Plaine d'Alsace. En effet, cette grille procure une évaluation d'une sévérité intermédiaire entre les grilles de cours d'eau d'altitude de taille plus modeste et les grilles de cours d'eau de plaine.

Comme les 2 TGCA précités, la Dordogne est un hydrosystème déjà de grande taille, mais directement adossé à un bassin versant montagnard. D'autre part, ces 3 hydrosystèmes ont le point commun de s'écouler à l'aval de systèmes lénitiques de volume important qu'ils ont traversé (Léman pour le Rhône, Lac de Constance pour le Rhin, grosses retenues hydroélectriques comme Bort-les orgues, Entraygues sur Truyère pour l'hydrosystème Dordogne + affluents).

Dans le cadre du 1^{er} dispositif national proposé à l'intercalibration, l'envoi au GIG VLRs du jeu de donnée national pour l'intercalibration effectué en décembre 2015 s'est donc appuyé sur des calculs d'EQRs correspondant à cette Grille 6 Diatomées pour la Dordogne dans son début de parcours TGCE (HER 11), puis à la Grille 4 à partir de son entrée dans l'HER 14.

5.4.4. Résumé de l'exercice d'intercalibration N° 1 du X-GIG « Large Rivers »

a) *Eléments de contexte général (Intercalibration Européenne) :*

L'exercice d'intercalibration « Large Rivers » auquel nous avons participé en 2015-2016, intitulé « Fit-in Exercise », visait à comparer les résultats d'évaluation procurés par les Pays-membres participant à ce second exercice au niveau d'évaluation adopté au cours d'un 1^{er} exercice d'intercalibration conduit précédemment par le X-GIG Large Rivers (2011-2012).

En effet, dans le cadre des procédures européennes d'intercalibration, ce sont les résultats du 1^{er} exercice collectif réalisé et validé par l'U.E. sur un type donné d'hydrosystème (ici, les TGCE de plus de 10 000 km² de surface cumulée de BV au site d'observation) qui servent ensuite de référence pour inter-comparer les systèmes d'évaluation récemment mis au point et proposés à l'inter-calibration par de nouveaux Pays-Membres. Ce principe de recours permanent au 1^{er} exercice validé sur le même type européen de cours d'eau permet d'assurer l'homogénéité, dans le temps et au travers d'exercices menés sur la base de jeux de données différents, du niveau d'évaluation faisant référence pour les inter-comparaisons réalisées au niveau Européen et pour la validation de leurs résultats.

b) *Résumé du 1^{er} exercice réalisé par le X-GIG Large Rivers (2011-2012) :*

Au niveau Européen, le premier exercice d'intercalibration mené sur ce type de cours d'eau « large Rivers » (bassin versant intégré > 10 000 km²) a concerné d'une part le maillon des macro-invertébrés benthiques, et d'autre part les diatomées benthiques. Il a été réalisé en 2011-2012 dans le cadre du X-GIG Large Rivers, sous la responsabilité de 3 responsables de nationalité allemande, à savoir Franz Schöll (coordinateur principal de l'exercice), Sebastian Birk et Jürgen Böhmer.

Cette intercalibration s'est basée sur la réponse d'un étalon commun européen d'intercalibration diatomique, le diatom-ICM ou d-ICM (REF. 12 : Kelly et al, 2009) vis-à-vis d'un **gradient unique de PO4**. Cet étalon européen diatomique, déjà utilisé dans le cadre de plusieurs exercices précédents d'intercalibration diatomique menés sur des types de cours d'eau plus petits, a été construit sur la base de 2 indices diatomiques internationalement reconnus : l'IPS de Coste déjà cité (cf. REF. 6) et le Trophic Index de Rott (REF. 13 : Rott et al, 1999).

Pour le 1^{er} exercice X-GIG, le gradient d'orthophosphates, représentatif du processus d'eutrophisation des cours d'eau, a été repéré comme **le principal gradient d'anthropisation** ayant contribué à structurer le jeu de données d'intercalibration (constitué par l'agrégation des jeux de données nationaux des pays-membres participants). Les réponses des systèmes d'évaluation nationaux au gradient de PO4 et leur niveau d'évaluation ont donc été intercomparés au travers de la réponse de l'étalon commun d-ICM à ce gradient trophique.

Ce premier exercice a associé, au titre de participants principaux, les 12 Etats-Membres suivants : Allemagne, Autriche, Belgique (Wallonie), Belgique (Flandres), République Tchèque, Estonie, Finlande, Hongrie, Pays-Bas, Slovaquie, Slovénie, Suède.

Faute d'un recul suffisant sur ce sujet, la France a participé à l'exercice **au titre d'observateur externe** et en fournissant notamment un jeu de données national et, mais ne s'est pas faite inter-calibrer au final, et les données qu'elle a fourni pour ce 1^{er} Round n'ont pas servi à la définition de la vision partagée du niveau d'évaluation.

Pour cadrer la mise en application opérationnelle de la DCE selon un cadrage méthodologique adéquat, il a été mis en place sur le site UE CIRCA une liste de documents-guides (ou Guidances) édités par le JRC ou par différents groupes méthodologiques ayant travaillé pour le compte de l'UE dans le cadre de la Common Implementation Strategy (CIS), comme ECOSTAT, REFCOND etc..., accessibles au lien suivant :

http://ec.europa.eu/environment/water/water-framework/facts_figures/guidance_docs_en.htm

La réalisation des exercices d'intercalibration programmés au niveau de l'UE est ainsi accompagnée de documents-guides d'application des différents processus, dont le Document-Guide ECOSTAT N° 14 « Implementation Strategy » (REF. 14 : ECOSTAT Guidance 14, 2011) qui permet de décrire 3 grands groupes d'options d'intercalibration (et de nombreuses variantes) en fonction des conditions de réalisation de l'exercice (individuellement au niveau national, ou au sein d'un GIG) et des caractéristiques des jeux de données des pays participants.

Au cours de ce premier exercice, il est très vite devenu évident qu'il n'existait plus ou quasiment plus de situations de référence vraie, à l'échelle Européenne, sur ce type de grands cours d'eau, rendant impossible l'application de l'**Option 1** (niveau d'altération mesuré par **écart direct avec le niveau de référence vraie correspondant**). La méthodologie mise en place s'est donc positionnée dans le grand groupe de l'**Option 2**, à savoir d'une part que l'exercice d'intercalibration était réalisé **dans un cadre collectif à l'échelle d'un GIG** (ici le X-GIG Large Rivers), et faisait d'autre part recours à une **référence alternative commune pour l'interétalonnage** (common benchmark), à savoir un étalon commun d'intercalibration. Dans le cas des diatomées, cet étalon commun était le diatom-ICM (ou d-ICM) cité plus haut. Pour mémoire, l'**Option 3** concernerait pour sa part des exercices d'intercalibration qui seraient **menés de façon autonome** par un **Pays-Membre particulier**, par exemple en intercomparant sa propre méthode nationale au résultat déjà obtenu par un étalon commun européen (ICM) dans le cadre d'un exercice collectif réalisé antérieurement.

L'exercice collectif mené dans le cadre de ce 1^{er} X-GIG visait à inter-comparer directement les résultats des systèmes nationaux via la réponse de cet étalon commun. Cependant (**Figure 39**), outre l'absence complète de sites de référence vraie dans tout le jeu de données d'intercalibration Large Rivers disponible, il a été remarqué que l'ampleur des gradients présents au niveau de chaque pays participant était très sensiblement différente (certains pays avec des stations situées sur un seul grand continuum fluvial de type Rhin ou Danube n'ayant d'ailleurs pas de gradient de PO₄ réellement structuré et d'ampleur suffisante) et ne se recoupait pas suffisamment, sur la base du gradient de pression présent dans chaque pays, pour une comparaison commune associant en même temps les jeux de données de tous les Pays-Membres.

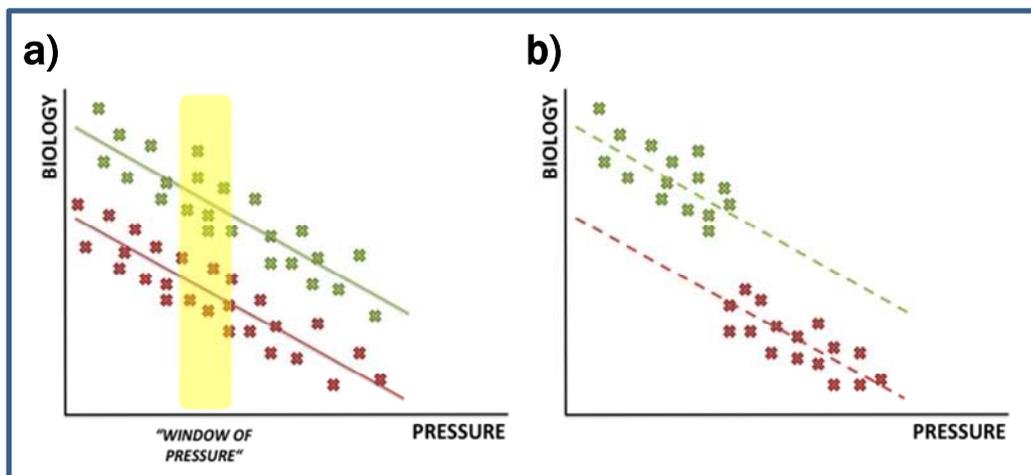


Figure 39 : les 2 approches fondamentales du benchmarking en intercalibration : a) approche se référant à des sites de référence et à un gradient de pression comparable ; b) » Approche de benchmarking continu (données des pays couvrant des gradients discontinus)

Il n'était donc pas possible dans ces conditions de faire recours à un niveau de référence vraie pour chaque jeu de données national, ni d'inter-étalonner les méthodes nationales à des endroits correspondant à des fenêtres de pression commune à tous les jeux de données nationaux, du type intercomparaison de différentes limites de classes de qualité (cf. sous-figure a, partie gauche de la **Figure 39**), par exemple limite BE-EM.

Par conséquent, c'est une **analyse comparative continue** sans intervention d'un niveau de référence vraie qui a été mise en œuvre (cf. sous-figure b, partie droite de la Figure 39). Elle s'est appuyée sur la combinaison des différents jeux de données nationaux et a été réalisée sur la base des concentrations d'orthophosphates (P-PO₄) trouvées dans les analyses d'eau.

Ce 1^{er} exercice a ainsi porté sur un nombre total de 410 doublets de données associant les moyennes annuelles d'orthophosphates aux données d'évaluation diatomique basées sur les relevés prélevés aux mêmes sites, comme indiqué dans le rapport final d'intercalibration X-GIG (REF. 15 : Schöll et al., 2012).

En résumé, la normalisation des méthodes d'évaluation nationale a donc été effectuée, non pas en fonction de niveaux de référence repérés par jeu de donnée national et via des petites plages présélectionnées de gradients, du type limite de classes de qualité High-Good (cf. partie a de la Figure 39), mais par un benchmarking continu prenant en compte l'ensemble du continuum (cf. Partie b de la Figure 39).

L'objectif général de la normalisation comparative en intercalibration est d'inter-comparer judicieusement les résultats des méthodes nationales d'évaluation de l'Etat Ecologique, ce qui implique d'avoir identifié et éliminé au préalable les différences inter-méthodes **qui ne sont pas d'origine anthropique**, mais qui génèrent des **décalages systématiques** liés à une méthodologie, une biogéographie, une typologie différentes, etc...

Si ces écarts systématiques non liés à l'altération anthropique sont ignorés, ils peuvent revêtir une influence déterminante qui viendra bruyter très sensiblement l'exercice de comparabilité. Le X-GIG a donc veillé à contrôler les effets de ce type de facteurs non-anthropiques d'influence sur les métriques communes afin de révéler et de prendre en compte les divergences parasites (*i.e.* non-liées au facteur d'altération anthropique pris en compte) à caractère systématique. Le gradient-cible de pression anthropique pris en compte dans cet exercice était celui des orthophosphates après transformation [Ln (P-PO₄)].

Pour l'obtention des valeurs de correction adéquates pour chaque jeu de données national, il a été utilisé un modèle linéaire à effets mixtes (cf. **Figure 40**, partie b). Dans le cadre de sa mise en œuvre, la métrique commune (le d-ICM) était la variable dépendante ; les variables de pression étaient un assortiment de covariables sélectionnées dont la fourniture était requise auprès de chaque Etat-Membre participant ; le pays (c'est-à-dire le jeu de données national) étant considéré comme un facteur fixe.

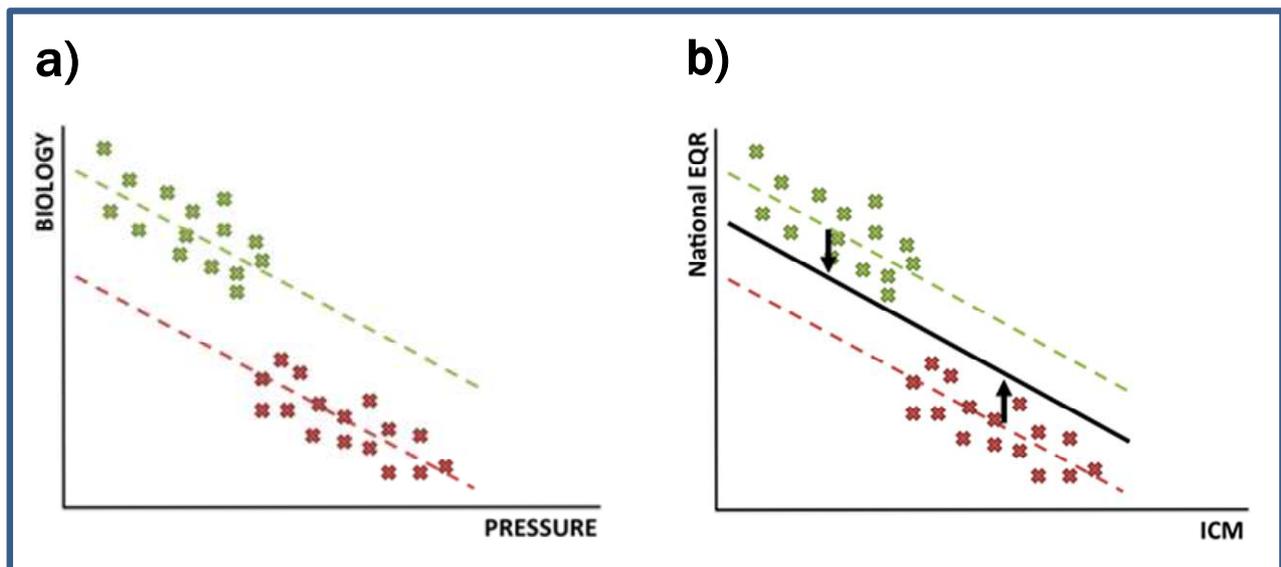


Figure 40 : Réajustement des EQR nationaux (rouge et vert) en fonction du décalage systématique mesuré par rapport à l'ajustement modélisé (en noir) de l'ICM (modèle linéaire à effets mixtes)

La **Figure 41** en page suivante présente l'organigramme de la procédure analytique déroulée par le X-GIG en vue de comparer les positions des limites de classe nationales de qualité, une fois re-calées par rapport à l'ajustement pression-impact modélisé par l'étalon européen (d-ICM).

Sans entrer dans tous les détails techniques, le modèle a fourni et quantifié les décalages systématiques moyens présentés par chaque jeu de données national par rapport à l'ajustement linéaire obtenu par l'ICM modélisé (ajustement obtenu à l'aide d'un modèle linéaire à effets mixtes). Ces décalages systématiques (« country-specific offsets » dans la figure) ont ensuite été corrigés pour re-caler le positionnement des EQR des pays par rapport au d-ICM modélisé (cf. partie droite de la Figure 40).

Le niveau d'évaluation par système national et par grand type de TGCE (faible alcalinité <0,5meq CO₃/l / versus alcalinité moyenne à forte >0,5meq CO₃/l) a ensuite été contrôlé, une fois ces biais systématiques réajustés, en établissant la bonne correspondance entre métrique nationale renormée en EQR et réponse du d-ICM modélisé, tant en terme de correspondance de limites de classes de qualité qu'en termes d'ajustement des points dans un graphique biplot par pays.

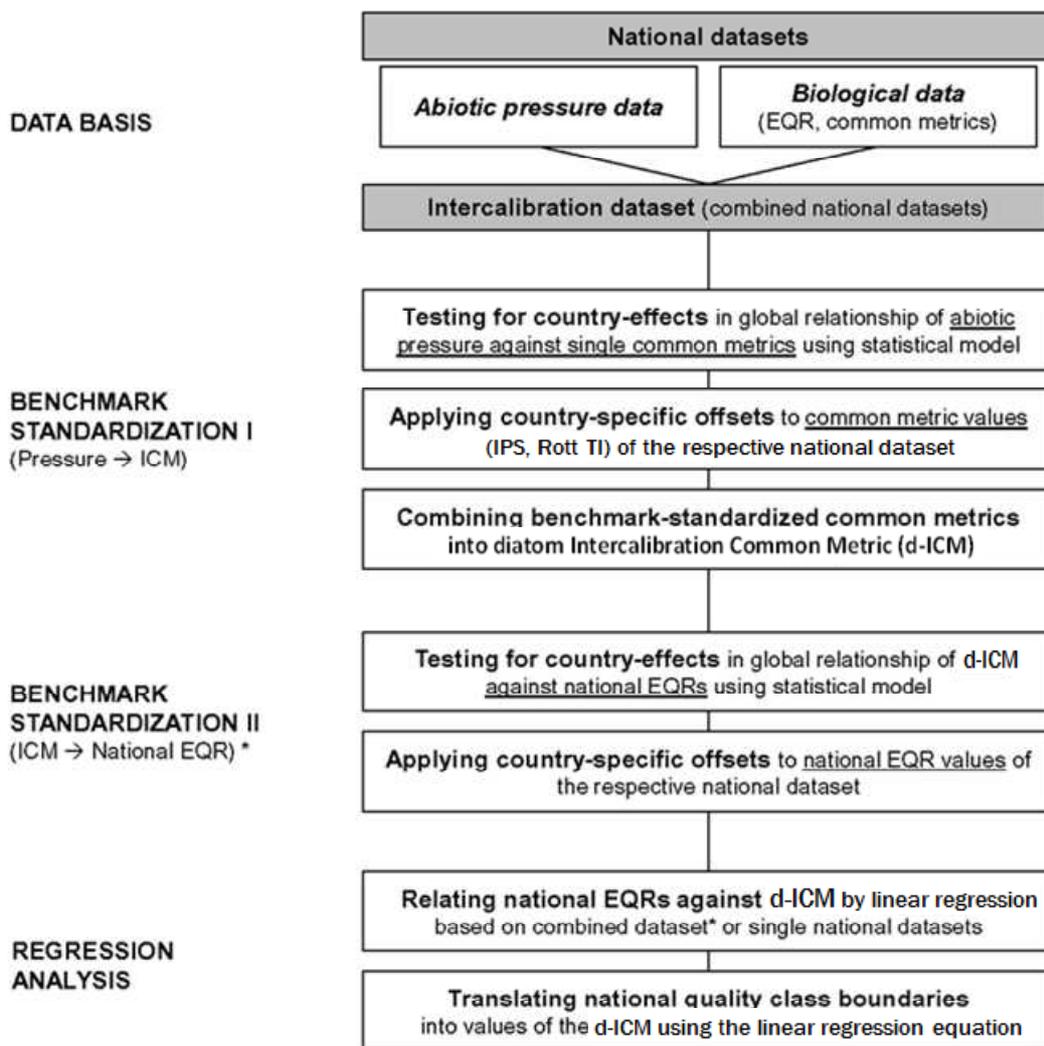


Figure 41 : Organigramme de la méthodologie suivie par le X-GIG pour le 1^{er} exercice d'intercalibration (2011-2012)

Les biais obtenus par rapport à l'ajustement global moyen utilisant toutes les données nationales du groupe (d-ICM modélisé) ont ensuite servi à agréer d'emblée, ou à refuser la DCE-compatibilité du système national par type dans son état actuel.

Suite aux problèmes de dépassements de biais, certains pays ont pu obtenir l'aval de leurs tutelles et réajuster rapidement leur niveau et leur future pratique nationale pour les types défectueux, selon les recommandations du GIG (pratique corrective suivie d'intercalibration OK).

D'autres qui ne pouvaient pas réagir assez vite, ou qui avaient un problème structurel d'assise insuffisante de données nationales (petits pays concernés par un seul TGCE Européen, comme le Rhin ou le Danube) pour établir un système d'évaluation autonome et bien calé, ont dû par force s'aligner sur le système d'un autre pays-membre voisin concerné par le même très grand hydrosystème et considéré OK dans le cadre de l'exercice.

Ce 1^{er} exercice d'intercalibration X-GIG s'est conclu par l'édition d'un rapport final d'intercalibration (cf. REF. 15) qui a été validé l'année suivante par l'U.E.

5.4.5. 1^{ère} itération de l'exercice d'intercalibration 2015-2016 « Fit-in exercise »

Les Etats-Membres participants à ce Fit-in Exercise, qui s'est mis en place à partir de Juin 2015, ont été au nombre de 8, à savoir : Bulgarie, coordonnateur de l'exercice en collaboration avec l'Autriche (qui avait participé au 1^{er} exercice et permettait d'assurer le lien avec celui-ci), Belgique (Flandres), Croatie, Espagne, France, Italie, Portugal, Roumanie.

Au niveau Français, les personnes de notre collectivité nationale impliquées dans le suivi de ce dossier ont été : François Delmas, Juliette Rosebery et Julie Guéguen (Irstea Bordeaux / UR EABX / Equipe de Recherche CARMA, Yorick Reyjol (ONEMA/ DAST, puis AFB), Marina Le Loarer, Stéphane Lucet et Jean-Pierre Cabaret (Ministère en charge de l'Environnement / Direction de l'Eau).

Après une phase de complétion échelonnée de Septembre à Décembre 2015, les caractéristiques quantitatives des jeux de données nationaux finalement fournis sont résumées dans le **Tableau 3** ci-dessous :

Tableau 4 : Jeux de données nationaux complets finalement fournis par les Etats-membres participants

MS	National types	Sites	Rivers	Biological data	Pressure data
Belgium	1	3	1	3	3
<i>Belgium - NL dataset*</i>	-	-	-	10*	10*
Bulgaria	1	15	1	23	23
Croatia	3	31	4	55	55
France	3	95	20	221	221
Italy	2	32	3	87	71
Portugal	1	20	1	20	19
Spain	1	48	5	75	74
Romania	4	23	6	158	158
<i>Romania - HU dataset*</i>	-	-	-	65*	65*

Les pays de ce groupe, selon leur taille géographique et les cours d'eau de grande taille qui y naissent ou qui ne font que les traverser, sont diversement concernés par le type Européen « Large Rivers ». A l'extrême, certains pays comme la Bulgarie et le Portugal amènent des données de plusieurs sites qui concernent tous un seul et même cours d'eau, et ne présentent donc pas forcément un gradient de pression d'anthropisation suffisamment étendu pour caler la performance de leur système national et même du d-ICM. Dans un tel contexte, la fourniture par d'autres Pays-Membres de données représentatives d'un gradient d'anthropisation plus étendu sont appréciables pour l'intercalibration de ces pays particuliers et pour le meilleur déroulement de l'exercice lui-même.

Pour réaliser cet exercice, 5 pays seulement ont pu fournir plus de 60 données. Quelques pays dotés d'une plus grande diversité de TGCE et de situations vis-à-vis du gradient d'anthropisation, dont la France, ont pu fournir les jeux de données les plus substantiels. La France présentait la particularité d'être concernée sur son territoire national par des TGCE diversifiés, représentatifs à la fois de conditions d'aval immédiat de montagnes, de plaines carbonatées et de zones méditerranéennes, et un cours d'eau peu minéralisé (la Vilaine aval dans l'HER 13 « Massif Armoricaïn). Après une re-sélection quantitative substantielle (un peu plus d'1/3 des données mobilisables), elle a au final contribué le plus fortement (221 relevés nationaux) à la fourniture de ce jeu de données d'intercalibration.

Dans ce second exercice d'intercalibration, la faisabilité initiale du processus a été confirmée pour tous les États membres participants en termes de typologie, de pression anthropique-cible (PO4), de concepts d'évaluation, de fourniture des différents types de descripteurs demandés dans le cadre de l'exercice, en association avec les relevés diatomiques nationaux. Toutes les méthodes nationales proposées à l'intercalibration se sont révélées conformes à la DCE et les ensembles de données nationaux ont rempli les critères d'acceptation des données à tous égards. Cependant, de la même manière que pour l'exercice X-GIG précédent, la gamme de gradients de pression couverts par les pays adhérents est apparue plus limitée sur ces grands systèmes que pour des cours d'eau plus petits.

Du fait du nombre de TGCE présents sur son territoire, qui couvrent une large palette d'environnements naturels, le jeu de données fourni par la France était le plus large, soit 221 relevés représentant 95 sites différents de réseaux (soit plus de 30% du jeu de données complet de ce 2^{ème} exercice d'intercalibration, qui représente un total de 717 relevés). Par rapport au gradient de pollution considéré (PO4), la plupart des autres jeux de données nationaux couvrent des gradients de réponse diatomique relativement courts s'étendant au maximum sur 3 classes de qualité de l'eau parmi les 5 classes classiquement utilisée dans le cadre d'application de la DCE (de H = High à B = Bad).

La partie gauche de la **Figure 42** ci-dessous présente les caractéristiques de distribution de la pression de trophie (PO4) sur 5 jeux de données de pays-membres participants dans l'état où ils ont été fournis fin 2015 (les jeux de données de certains pays comme l'Espagne n'étant pas encore complets, et d'autres pays pouvant n'être concernés que par un seul grand à très grand fleuve et en conséquence, ne présenter qu'un gradient relativement étroit de pression anthropique). Ces box-plots sont confrontés au gradient présent sur tout le jeu de donnée d'intercalibration précédent X-GIG (à droite du trait bleu discontinu).

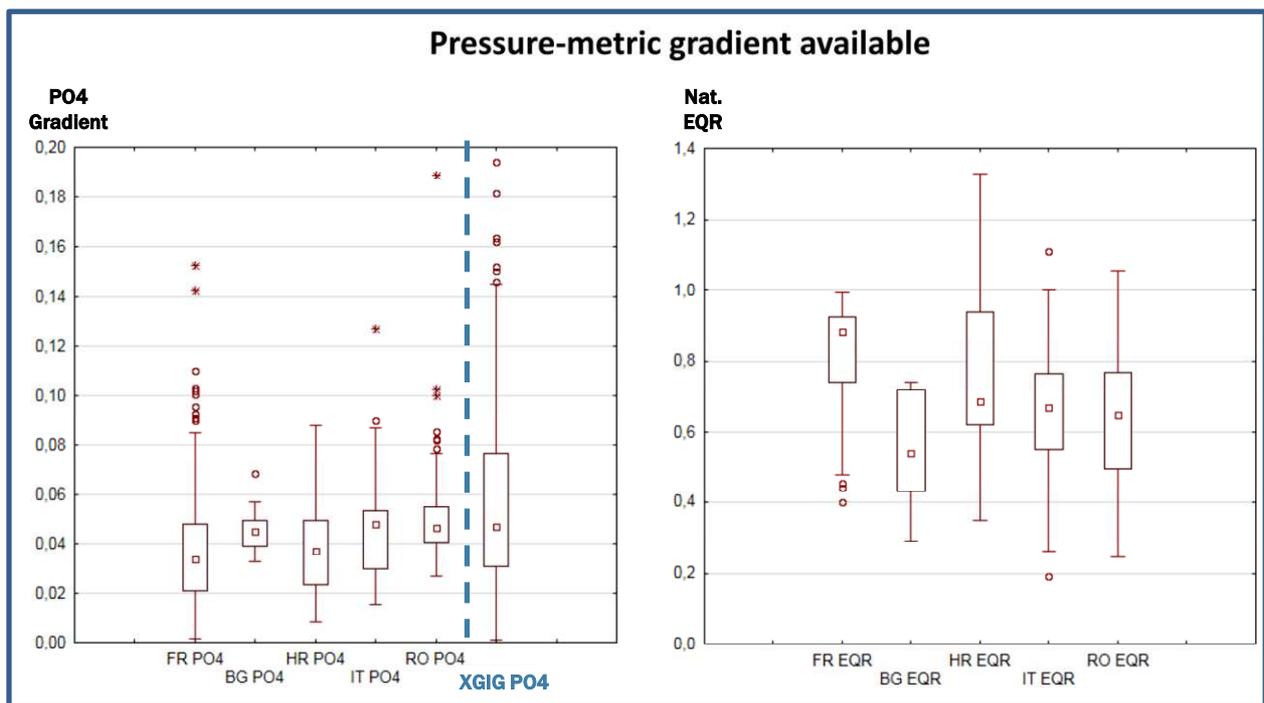


Figure 42 : Gradients de PO₄ transformés Ln (à gauche) et réponses diatomiques en EQR nationaux (à droite) présentés par les jeux de données du Fit-in Exercice « large Rivers » 2015 – 2016

On peut remarquer que, si le niveau moyen des concentrations moyennes annuelles en PO₄, après transformation Log (népérien), est assez proche à l'échelle globale des 2 exercices d'intercomparaison (cf. carrés au centre des box-plots par pays, d'un niveau assez comparable à celui de tout le jeu de données du précédent X-GIG), le précédent exercice X-GIG présentait un niveau de pression légèrement plus élevé des percentiles 75 et 90, donc une assise de distribution un peu plus large du gradient de PO₄ au niveau des boîtes à moustaches que l'actuel Fit-in Exercice. Cependant, en prenant aussi en compte les outliers dans l'amplitude des concentrations représentée, la couverture des possibles dans les 2 exercices est en fait du même ordre.

La partie droite de la Figure 42 présente la distribution des valeurs d'EQR des métriques diatomiques des mêmes 5 pays-membres participants dans leur système d'évaluation national. A ce stade encore très préliminaire, on peut déjà remarquer 2 faits susceptibles d'influer de façon assez lourde sur le processus d'intercalibration, qui vont nécessiter une vigilance particulière :

- 1) Les principes de construction des systèmes nationaux et les bases de calcul d'EQR semblent différer très sensiblement selon les pays, certains d'entre eux pouvant par construction arriver à des valeurs très supérieures à 1 (ici, la Hongrie approchant 1,4 et l'Espagne, non-représentée pour raison de jeu de données encore incomplet à cette époque, pouvant même atteindre 1,5). Ces différences de construction laissent craindre un principe d'évaluation basé sur un niveau de référence trop bas (puisque des sites à évaluer, qui ne sont pas des sites de référence (ce statut de référence vraie n'existant plus en Europe pour de tels grands hydrosystèmes) peuvent atteindre 1,5 fois le niveau de référence ou de référence par défaut adopté par un pays pour évaluer ce type de TGCE).

Or, au sens de la DCE, un écart mesuré par rapport au niveau moyen de référence (établi par construction à la valeur de 1), qui produirait une valeur de + 1,5 sur certains relevés de certains pays (soit 150 % du niveau de référence) devient complètement ininterprétable sur le plan de la signification biologique et de l'état écologique. L'exercice d'intercalibration se faisant par rapport aux valeurs d'EQRs nationaux régionalisés et d'autres pays semblant évaluer leurs cours d'eau sur une base plus réaliste mais aussi, par construction, plus sévère (puisque la valeur maximale atteinte sur leurs sites dans le meilleur état ne dépasse pas, ou ne dépasse que très occasionnellement et très modérément la valeur moyenne de référence de 1), ce premier problème est susceptible de créer des artefacts de calculs de pentes entre les pays-membres participants.

2) Les différents Pays-Membres ne sont probablement pas confrontés aux mêmes gradients et aux mêmes qualités de fleuves selon leur territoire et les pressions anthropiques qu'ils portent. Cependant, la différence très sensible de calage de niveau de leurs box-plots respectifs est très probablement en partie (à vérifier...) la conséquence de décalages systématiques de niveau d'évaluation occasionnés par un calage différent de leurs systèmes nationaux d'évaluation (cf. Figure 40 relative aux biais systématiques et à la nécessité d'intervention d'un principe de leur évaluation et de leur correction). Aux 2 extrêmes, on peut remarquer une évaluation en apparence plus sévère en moyenne de la Bulgarie et probablement plus laxiste en moyenne de la Hongrie, ainsi que de l'Espagne (non intégrée dans cette figure). Un fait rassurant est que la France et l'Italie, qui ont à coup sûr des caractéristiques assez proches des conditions naturelles et de la biogéographie de leurs grands hydrosystèmes, ainsi que la Roumanie (contexte local moins bien connu de nous), semblent évaluer leurs cours d'eau dans leur système national de façon assez comparable, tant au niveau de leur construction d'EQR par rapport à un niveau de référence (valeurs plafonnées à 1 ou pouvant dépasser légèrement cette valeur dans de rares cas) que du niveau moyen et de l'amplitude de leur gamme d'évaluation. Pour rappel, la France et l'Italie présentent un assortiment varié de TGCE dont certains arrivent directement du domaine montagnard et d'autres sont plus typiques d'un contexte minéralisé de faible altitude, voire même du domaine méditerranéen.

Pour la réalisation de l'exercice commun d'intercalibration, le jeu de données d'intercalibration constitué par l'agrégation des jeux de données fournis par les Pays-Membres se caractérisait par des gradients de pression limités et pouvant être discontinus selon le Pays-Membre, et par l'absence de références ou de références alternatives internes au jeu de données.

Dans le cadre des procédures prévues pour adapter les classifications nationales nouvelles ou révisées aux résultats d'un exercice d'intercalibration précédent validé (dans notre cas, le 1^{er} exercice réalisé par le X-GIG), la coordinatrice a décidé d'adopter une **analyse comparative continue** s'appuyant sur une normalisation des métriques. Parmi les diverses possibilités envisageables, définies et décrites dans le Document-Guide N° 30 de la Common Implementation Strategy (**REF. 15** : Fit-in Guidance Nr 30, 2015), le choix méthodologique finalement réalisé correspond à l'**Option 2 (cas d'ajustement A2, avec analyse comparative continue)**, en utilisant une comparaison indirecte via une métrique d'inter-étalonnage commune (ici, le diatom-ICM).

Ce cas de figure, qui correspond à la procédure 5.2 de ce guide méthodologique, doit être appliqué : **1)** lorsqu'il n'existe pas de situations de référence vraie dans le cadre de l'exercice à mener, mais **2)** qu'il existe un étalon commun d'intercalibration validé (c'est le cas ici avec le d-ICM...) ; enfin **3)**, lorsqu'un exercice d'intercalibration précédent a été réalisé et validé, mais que la plage du gradient de pression disponible dans les jeux de données fournis par les différents États membres ne se chevauche pas suffisamment, ce qui empêche l'utilisation de points de référence alternatifs communs à tous les États-Membres au long du gradient. Dans un tel cas, la relation globale entre la métrique commune et le gradient de pression présenté dans le cadre du précédent exercice réalisé (ici, celui du X-GIG) doit servir de base à l'analyse comparative.

Les conditions de mise en œuvre du cas d'ajustement A2 ont été les suivantes :

L'étalon de référence européen qui a été appliqué dans le cadre du précédent X-GIG 2011-2012, ainsi que dans le cadre du « Fit-in Exercise » qui nous a concernés en 2016, tel que décrit par Kelly et al, 2009 (cf. **REF. 12**), est le diatom-ICM ou d-ICM. Comme déjà décrit antérieurement, il est composé de 2 métriques diatomiques largement reconnues et faisant référence au niveau Européen, à savoir :

- L'**IPS**, ou Indice de Polluo-Sensibilité (Coste in CEMAGREF, 1982) (cf. **REF. 6**)
- Le **RTI**, ou Indice Trophique de Rott (Rott *et al.* 1999) (cf. **REF. 13**).

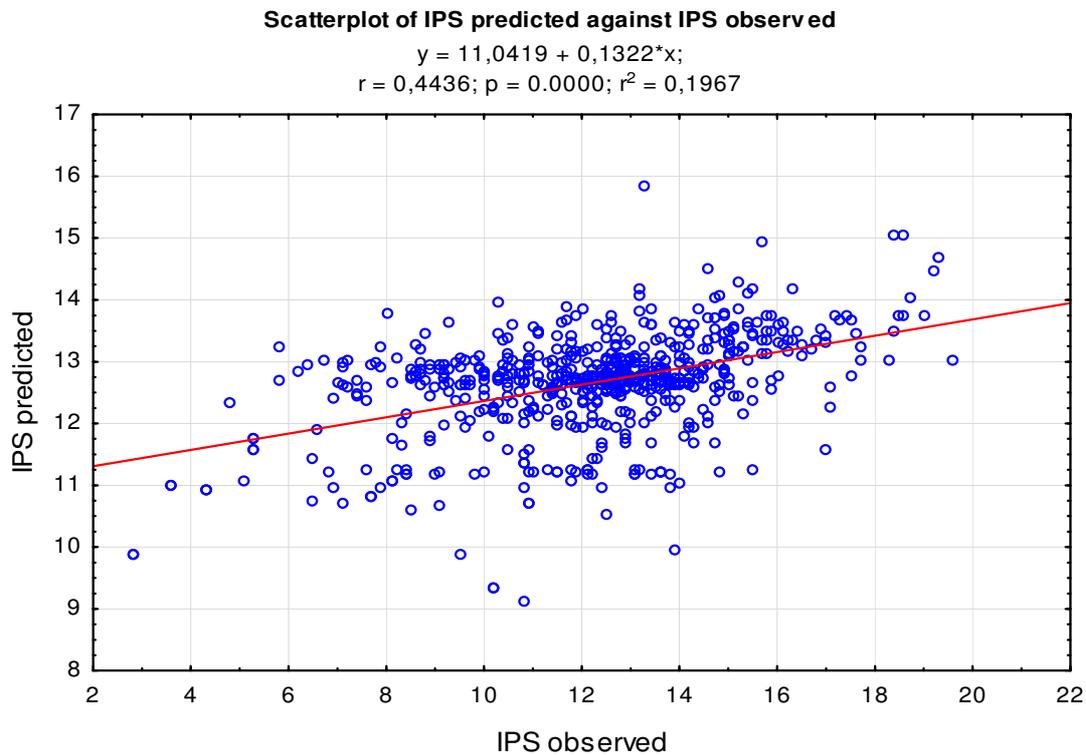
Le d-ICM tel que décrit ci-dessus, déjà utilisé dans l'exercice précédent du X-GIG "large Rivers", établit la moyenne normalisée de référence entre les 2 métriques diatomiques qui le composent (IPS et RTI), en fonction du gradient de PO4 rencontré. En l'absence de situations de référence valides dans les jeux de données nationaux « Large Rivers », ces 2 métriques diatomiques ont été re-normalisées par rapport au précédent exercice en utilisant le percentile 10 et le percentile 90 des valeurs obtenues dans les 2 cas à partir des jeux de données nationaux combinés.

Ainsi, l'ICM modélisé de référence, ou ICM-bm, correspond à la valeur de d-ICM qui serait obtenue dans l'ajustement linéaire « pression de PO4 - impact sur le d-ICM modélisé pour le précédent jeu de données X-GIG » (même ajustement percentile 10 - percentile 90), sur la base des valeurs d'EQR national obtenu par relevé national correspondant.

Pour le calcul des métriques et l'harmonisation interne des listes de taxons des pays-membres participant à l'exercice, la base de données originale complète de Martyn KELLY utilisée pour l'inter-étalonnage du 1^{er} exercice du XGIG « large Rivers » a été appliquée, y compris les scores par espèce, les poids métriques et les synonymes pour l'harmonisation taxonomique.

Démarche suivie :

a) Les 2 métriques constitutives du diatom-ICM sont tout d'abord séparément re-normées en fonction de l'ajustement qu'elles auraient présenté avec les concentrations de phosphore dans le 1^{er} jeu d'intercalibration X-GIG, produisant une métrique benchmark modélisée. (cf. **Figure 43** : cas de l'IPS. La même procédure est réalisée en parallèle avec le RTI).



*Figure 43 : Relation entre IPS observé et IPS benchmark prédit
(en fonction de la relation avec le PO4 pré-établie dans le cadre du précédent X-GIG "Large Rivers")*

L'ICM-benchmark modélisé (ou ICM-bm) est ensuite calculé à partir des deux métriques constitutives prédites selon la formule suivante :

$$\text{ICM-bm} = (\text{IPS_bm_norm} + \text{RTI_bm_norm})/2$$

b) Les points d'ancrage pour la normalisation des métriques (percentile 10-percentile 90), qui s'appuient sur les résultats de l'exercice X-GIG précédemment réalisé, sont les suivants :

- **15.64118 / 10.23263 pour l'IPS (sur une échelle de notation de 0 à 20),**
- **3.20493 / 2.41768 pour le RTI (sur une échelle de notation de 4 à 0)**

c) L'équation ordinaire de régression linéaire basée sur les moindres carrés (ou "OLS regression") du modèle global a permis de mettre en relation la métrique commune et la pression exercée dans l'exercice terminé. Ces informations sont nécessaires même si, à la fin de l'exercice, il a été conclu qu'une normalisation de référence n'était pas nécessaire (du fait que, dans le modèle associant l'ICM à la pression PO₄, l'effet du facteur fixe "Pays" ne s'est pas révélé significatif).

Relation des 2 métriques constitutives modélisées avec les teneurs en PO₄ :

- $IPS = -0.9183 * \ln_PO4 + 9.8987$
- $RTI = 0.24452 * \ln_PO4 + 3.63873$

d) Valeurs de la vision moyenne globale des limites H-G (High-Good) et G-M (Good-Moderate) sur l'échelle de la métrique commune modélisée (ICM-bm), obtenue à partir de l'exercice X-GIG précédent.

- **HG 0.603 - GM 0.312 pour les TGCE d'alcalinité moyenne à forte (> 0,5 meq CO3/l)**
- **HG 0.985 - GM 0.642 pour les TGCE à faible alcalinité (< 0,5 meq CO3/l)**

En date du 04 Mai 2016, la coordinatrice de l'exercice a fait circuler aux participants un premier draft de rapport d'intercalibration du 03-05-2016, sollicitant des Pays-Membres des réactions, observations, remarques et questions sur le contenu et les premiers résultats obtenus, suivi de l'envoi d'une 2ème version datée du 30-05-2016.

Ces 2 versions successives de draft de rapport d'intercalibration ont généré des échanges fournis dans le mois et demi qui a suivi. En effet, du point de vue de la France, différents aspects posaient question, ou ont nécessité des explications, ou ont justifié des remarques et suggestions vis-à-vis des modalités plus précises de la procédure suivie, justifiant l'envoi successif de 3 notes nationales techniques et d'expertise :

- la 1^{ère} du 08-05-2016 (**REF. 16** : Delmas, F., 2016-1) ; la 2^{ème} du 18-05-2016 (**REF. 17** : Delmas, F., 2016-2) et la 3^{ème} du 03-06 -2016 (**REF. 18** : Guéguen, J. & Delmas, F., 2016), en réaction aux différents fichiers de résultats d'étape envoyés séparément, ainsi qu'aux rédactions, puis modifications incluses dans les 2 versions successives de draft de rapport d'intercalibration envoyées par la coordinatrice pour avis et remarques.

Tout d'abord, sur le plan de la réalisation globale de l'exercice et entre les pays participants, l'exercice de Fit-in présentait des ajustements linéaires de pentes relativement homogènes entre les EQR nationaux et l'ICM-bm modélisé mais, de façon peu compréhensible, très différentes par rapport à l'ajustement linéaire obtenu dans le cadre du 1^{er} exercice X-GIG.

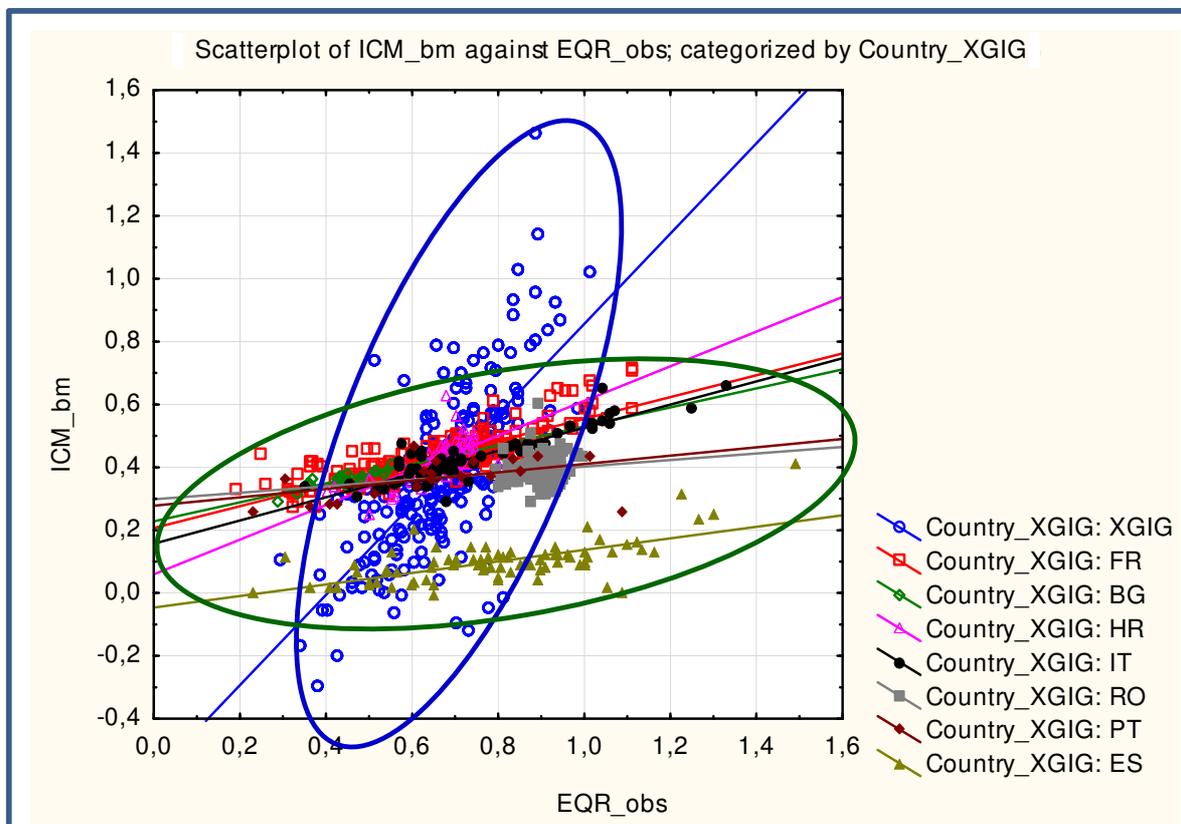


Figure 44 : 1^{ère} mise en relation entre EQRs nationaux d'évaluation diatomique du 2nd Fit-in exercice, du jeu de données du 1^{er} X-GIG exercice (points et ajustement en bleu), avec l'ICM-bm modélisé.

Cet écart très sensible de relation-type apparaissait injustifié par rapport à la partie gauche de la Figure 42, qui montrait des gradients de phosphore d'amplitude plutôt comparable dans les 2 exercices. D'autre part, certains des fleuves concernés par le 1^{er} et le 2^{ème} exercice étaient communs (le Rhin, le Danube, etc...), rendant plutôt suspect ce fort écart d'ajustement du même d-ICM vis-à-vis du gradient de phosphore.

Dans le cadre des envois successifs de notes techniques nationales précitées, nous avons donc fait remonter des remarques suggérant un problème de différence méthodologique intervenue entre le 1^{er} et le 2^{ème} exercice dans la mise en relation des concentrations de PO4 et l'ICM-bm. 3 pistes susceptibles de provoquer de tels écarts entre les 2 exercices ont été suggérées à la coordinatrice, à savoir :

- une unité non-homogène d'expression des concentrations de phosphore entre ces 2 exercices (PO4 vs P-PO4) ;
- 2) une procédure non-homogène pour effectuer la transformation log des données de concentrations en phosphore (l'écriture mathématique de la transformation et l'écriture de programmation sous R n'étant pas similaires, donc susceptibles de générer une différence d'échelle utilisée entre log népérien et log décimal) ;
- 3) un éventuel problème de seuils de quantification très différents entre les Pays-Membres et du mode de traitement des valeurs non-quantifiées pouvant occasionner, avec une formule de transformation du type $y = \log x$, une influence importante sur la pente de la relation.

Concernant cette 3^{ème} éventualité, en cas de transformation logarithmique servant ensuite à modéliser une relation pression-impact, nous aurions plutôt appliqué une formule du type $y = \log(x+1)$, qui conduit à une valeur minimale possible de 0 à l'origine de l'axe des y, et pas à une valeur minimale pouvant s'étendre plus ou moins loin dans le domaine à gauche de l'axe des y en fonction du nombre de chiffres après la virgule résultant du résultat direct d'analyse ou de calcul, arrondi ou non, et tendre vers "– l'infini". Ainsi, la formule de transformation log utilisée par le GIG peut conduire à des pentes sensiblement influencées par des artefacts, selon la façon de traiter les très faibles valeurs et/ou les valeurs non-quantifiées de PO4 ramenées, soit à zéro, ce qui conduit à une impossibilité ($\log [0] = - \text{l'infini}$), soit à une fraction du seuil de quantification évitant la valeur 0 (par exemple, 1/3 ou 1/2 du SQ). En d'autres termes, le calcul de moyenne annuelle réalisé par chaque Pays-Membre sur des valeurs de seuils de quantification différents, et qui conduisent à des valeurs basses de moyennes annuelles générées ou post-traitées de façons si différentes, va amplifier et structurer artificiellement un gradient de très faibles valeurs qui, à l'issue de la transformation, vont largement s'étendre vers le domaine négatif et être génératrices de pentes différentes, alors que ces très faibles valeurs juste au seuil de quantification du laboratoire ou en-dessous sont très peu représentatives d'un gradient réel d'influence du PO4 sur le compartiment diatomique.

N'ayant pas obtenu en retour de réponse claire sur l'origine réelle de ce problème étrange de très grande différence de pente, sur son origine et sur son mode de résolution, et du fait de la disparition du graphique de la Figure 44 dans les versions ultérieures du rapport d'intercalibration, il n'est donc pas complètement certain qu'il y ait été remédié pour les évolutions suivantes du processus d'intercalibration (cf. chapitre suivant).

Un autre problème rencontré lors de ce 1^{er} round de simulations, qui n'a vraiment pu être caractérisé de façon certaine qu'à partir du moment où la coordinatrice nous a fait parvenir ses tableaux détaillés de calcul (en date du 21-05-2016), était que la partie Française ne parvenait pas à reconnaître certaines valeurs de données présentées dans des biplots de mise en relation. Ce n'est que tardivement que nous avons pu acquérir la certitude que, malgré la réception d'un mail accusant bonne réception de ce fichier définitif, le jeu national utilisé pour les simulations dans l'exercice collectif n'était pas le dernier jeu de données complet et validé envoyé par la France le 12-12-2015, mais le jeu de données provisoire antérieur fourni sous délai contraint par la France le 15 Septembre 2015, en vue de sauvegarder sa participation effective à l'exercice.

Les résultats du 1^{er} draft de rapport d'intercalibration du 03-05-2016 incluaient donc quelques erreurs pour la mise en relation, les équations et les biais indiqués pour la France, qui ont été corrigés ensuite. A ces erreurs près, qui ne portaient que sur certains points Français et ne changeaient donc qu'à la marge la représentation du positionnement respectif des jeux de données des pays-membres, l'examen de la Figure 44 montre, comme déjà dit, des pentes relativement similaires de relation entre tous les Pays-Membres participants, signant de façon indirecte que leurs relations pression-impact entre phosphore et outil diatomique se présentent de façon assez homogène et cohérente au sein de l'exercice collectif qui nous a concernés.

Cependant, il est visible aussi que les Pays-Membres participants n'ont pas construit leurs EQR nationaux sur la même base, c'est-à-dire par rapport à un niveau de référence par défaut homogène entre pays. Du fait que les comparaisons sont effectuées en EQR, ces différences de principe de construction conduisent donc à des systèmes calés globalement sur des bases plus ou moins laxistes ou sévères (cf. les différences de niveau d'évaluation par exemple entre Espagne, pays dont le dispositif serait *a priori* le plus laxiste, et la France et l'Italie qui semblent présenter des dispositifs globalement plus sévères).

Cette différence de principe de construction du niveau de référence et des calculs d'EQRs pose un problème de comparabilité entre dispositifs du fait que certaines valeurs d'EQR exprimées dans certains systèmes nationaux d'évaluation montent jusqu'à une valeur de 1,4, voire 1,6 (le niveau moyen de référence par défaut dans ce même pays étant centré sur 1) ce qui, écologiquement parlant, ne correspond plus à rien en matière d'expression d'écart et de mesure du degré d'altération par rapport à un niveau de référence. Un tel dépassement du niveau moyen de référence (forcément de référence par défaut, dans le cas des TGCE) est probablement le signe d'un niveau de référence calé trop bas pour certains types nationaux de TGCE de Pays-Membres participants (ce problème concernant tout particulièrement l'Espagne).

Concernant la France, par construction, il est normal que quelques valeurs de sites en TBE dépassent modérément la valeur de 1, puisque le niveau de référence par défaut de 1 correspond à la **moyenne** des situations de référence par défaut (certaines étant un peu inférieures et d'autres étant un peu plus élevées que la valeur moyenne de 1). Dans cette figure, où on détecte des signes de l'erreur de version de notre jeu de données national, 2 relevés sont représentés à une valeur de 1,12 (en fait, leur vraie valeur est de 1,075), ce qui constitue un référentiel logique et qui, en tout cas, ne présente pas les mêmes problèmes de construction que pour certains autres pays participants.

L'ennui est qu'une inter-comparaison sur base d'EQR nationaux construits sur une base aussi hétérogène induit forcément des distorsions de pentes qui risquent d'influer, en fin d'exercice, sur l'aspect judicieux des écarts d'évaluation par rapport à l'ajustement obtenu dans l'exercice précédent. Aussi, nos 2 premières notes techniques nationales du 08-05-2016 et du 18-05-2016 ont fait état de ces constats auprès de la coordinatrice et soulevé les possibles répercussions des problèmes détectés sur les résultats de l'exercice.

Suite à la détection certaine du problème de version non-updatée de notre jeu de données national, signalée dans le cadre de la 2^{ème} note technique nationale envoyée (16-05-2016), la coordinatrice de l'exercice a ré-exécuté, à partir du bon jeu de données Français, la procédure complète de modélisation d'ajustement linéaire entre EQR-IBD et l'ICM-benchmark modélisé, et a ainsi pu nous transmettre des résultats plus fiables. Pour cette intercomparaison, les valeurs d'évaluation française en notes d'IBD₂₀₀₇ ont été transformées en EQRs nationaux selon les 4 grilles d'évaluation TGCE présentées en Figure 34 (soit la grille de type 2 pour les TGCE peu minéralisés ; la grille 4 pour les TGCE minéralisés de plaine ; la grille 5 pour 2 TGCE directement issus de zones de haute altitude du massif Alpin, à savoir l'Isère et la Durance ; enfin, la nouvelle grille 6 destinée à l'évaluation les TGCE Alpains plus éloignés des massifs montagneux, avec lacs et/ou gros barrages interposés (Dordogne, Rhin et Rhône). Les pentes ont été calculées sur la base des percentiles 10 et 90 du précédent jeu d'intercalibration X-GIG et du jeu de données national intercomparé. L'intercomparaison du jeu de données Français (en rouge) par rapport au jeu de données du précédent X-GIG (en bleu) s'établit comme illustré en **Figure 45** ci-dessous.

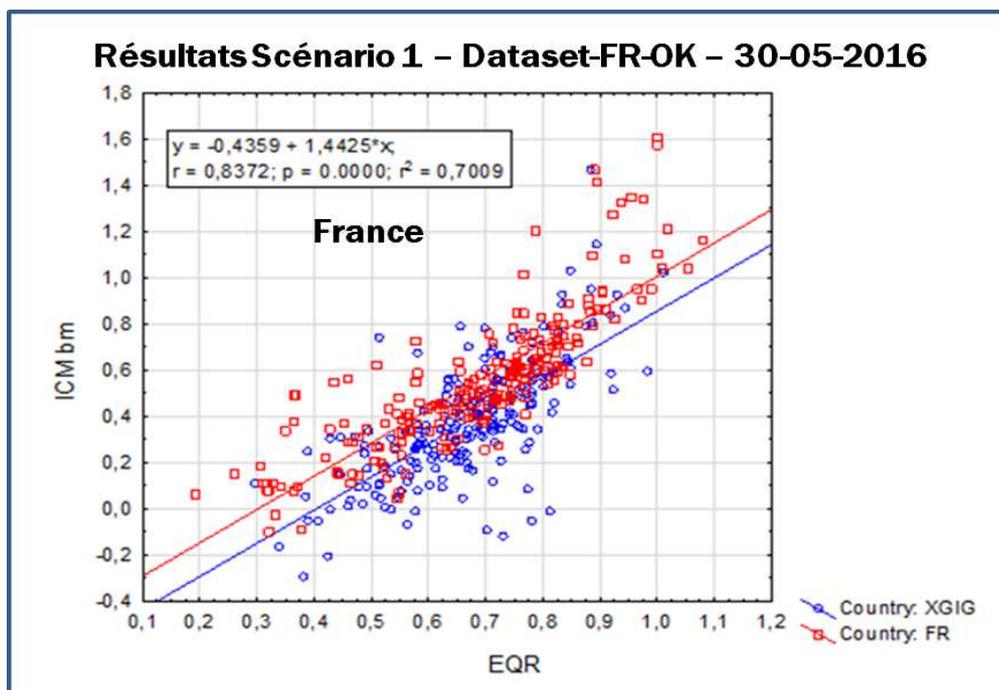


Figure 45 : Régression linéaire entre la méthode d'évaluation française (Scénario 1) et l'ICM-bm modélisé, en inter-comparaison avec le jeu de données du précédent X-GIG

A première vue, selon la méthodologie d'interprétation des résultats proposée dans la Guidance *ad hoc*, cette figure semblerait révéler que le 1^{er} prototype de système d'évaluation Français proposé est un peu plus sévère que la vision dominante du Bon Etat établie dans le cadre du précédent exercice X-GIG : cf. décalage vers le haut du nuage et de l'ajustement linéaire rouge (FR) par rapport à l'ajustement linéaire bleu (X-GIG).

Cependant, à titre d'éléments de modération de ce constat, on peut d'une part remarquer que, pour les 2 ajustements mais plus particulièrement pour l'ajustement du jeu de données Français (en rouge), la relation entre métrique nationale et ICM-bm modélisé ne semble pas obéir à un patron linéaire, mais plutôt curvilinéaire (cas de figure méthodologique d'inter-comparaison non prévu par les Guidances ECOSTAT). Cet aspect de déséquilibre des résidus aux 2 extrémités de la relation linéaire modélisée pour la France est matérialisé et aisément vérifiable à l'examen de la **Figure 46** ci-dessous. Des résidus équilibrés tout au long de la relation seraient le signe d'un ajustement linéaire judicieux, ce qui est loin d'être le cas ici. En effet, le présent type de répartition déséquilibrée des résidus illustre plutôt un patron de relation de type curvilinéaire. Et, dans un tel cas de figure, le principe d'ajustement et d'inter-comparaison de niveau d'évaluation entre le percentile 10 et 90 de chacune des métriques est susceptible à la fois de modifier la pente de l'ajustement linéaire obtenu entre les 2 métriques et d'engendrer un décalage avec l'ajustement du 1^{er} exercice dont l'origine pourrait être tout autant liée au déséquilibre des résidus à chaque extrémité de la relation qu'à une différence de niveau d'évaluation.

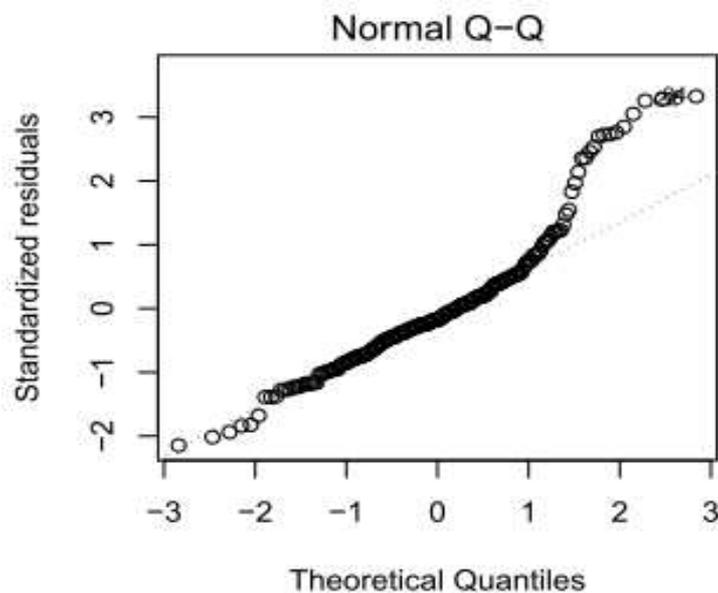


Figure 46 : Répartition des résidus en liaison avec l'ajustement linéaire produit pour la France en Figure 45

D'autre part, il est utile de garder en tête que la réponse de notre outil national d'évaluation basé sur l'IBD₂₀₀₇ n'est pas conditionnée uniquement par la teneur de l'eau en PO₄ (ici exprimée en moyenne annuelle). En effet, par construction, cet indice a pris en compte 7 différents types d'altérations intégrées sur une durée trimestrielle, dont certaines (Sat. O₂, NH₄, DBO₅, NKJ, NO₃ etc...) sont susceptibles d'intervenir souvent en mélange conjoint avec la première et, pour une bonne partie des relevés, d'**augmenter le niveau d'impact anthropique** mesuré par l'IBD sur les communautés diatomiques.

Ces altérations collatérales pouvant aussi présenter des différences d'intensité moyenne selon les pays, leurs particularités d'usages anthropiques et les pressions prises en compte dans le cadre de chaque outil national, il n'est donc pas anormal que l'IBD intègre pour partie l'effet de pressions non prises en compte dans cet exercice uniquement centré PO₄, et puisse procurer des résultats un peu plus sévères qu'attendu vis-à-vis de ce seul gradient.

Bien qu'il soit souhaitable d'évaluer les altérations à l'état écologique d'une façon la plus homogène possible avec les autres Pays-Membres, les 2 types d'informations de portée générale précitées, ainsi que d'autres plus spécifiques à certains contextes naturels particuliers de notre territoire, sont à garder en tête dans le présent cas de figure où notre dispositif d'évaluation serait à première vue jugé un peu plus sévère que la vision dominante résultant d'un exercice collectif.

La **Figure 47** en page suivante correspond à la même projection du jeu de données Français qu'en Figure 45. Mais en complément, elle matérialise à quels types nationaux appartiennent les relevés et les limites de classes d'évaluation selon les 2 systèmes, permettant d'examiner et d'interpréter certains aspects plus particuliers.

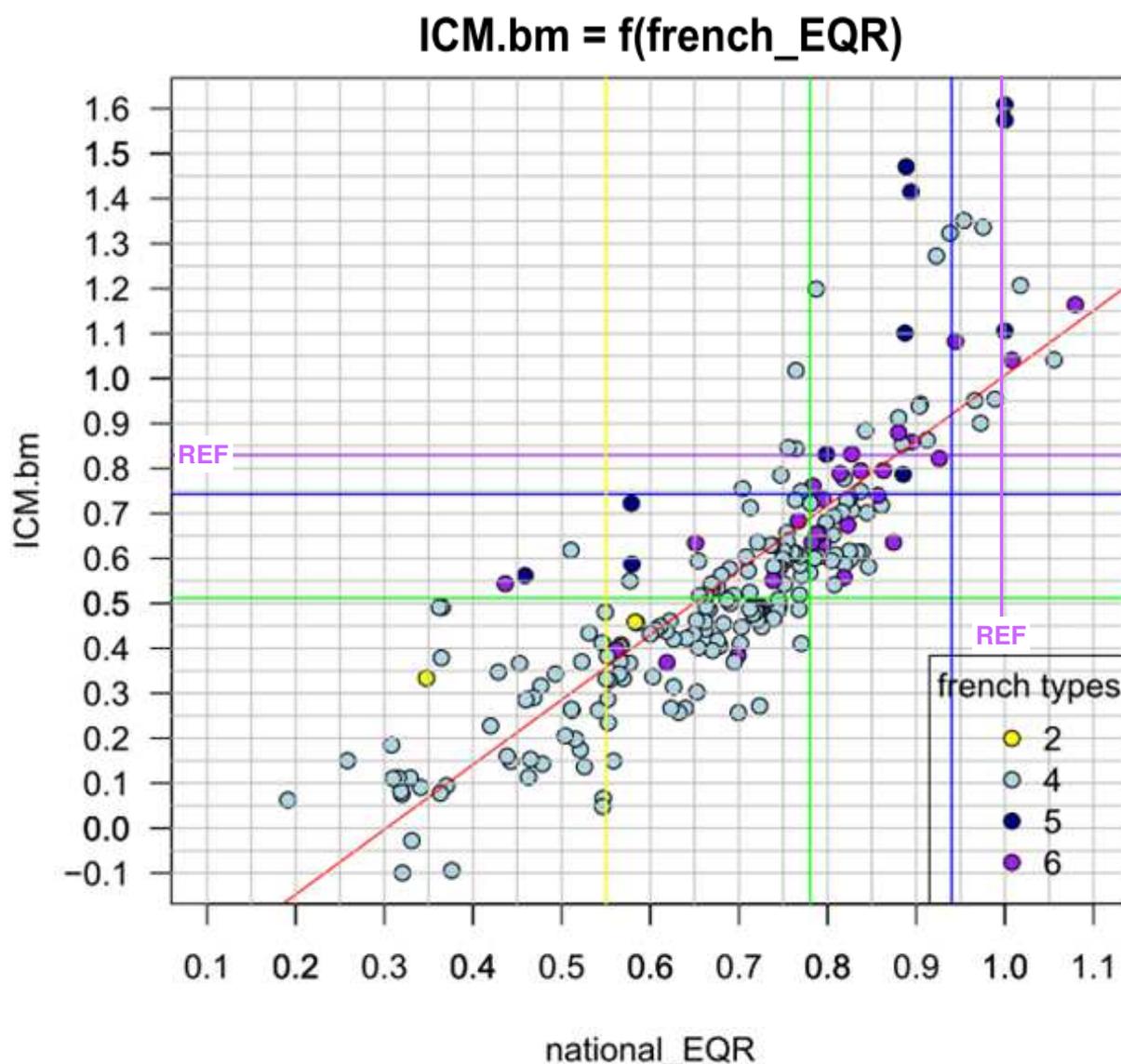


Figure 47 : Relation entre les EQR-IBD Français par type de TGCE et l'ICM-bm (1^{er} Scénario)

On peut tout d'abord remarquer que quelques prélèvements situés dans le coin en haut à droite de cette figure permettent d'obtenir de façon simultanée des scores élevés pour l'EQR-IBD et surtout, des scores exceptionnellement élevés (presque le double du niveau de référence pré-établi pour l'ICM-bm).

Ces scores d'ICM-bm étant uniquement déterminés à partir de la composition spécifique des assemblages de diatomées, le positionnement spécifique de ces points doit ici être interprété uniquement comme une très bonne évaluation biotique convergente d'entre les deux métriques biologiques (IBD-EQR et ICM-bm). Le décalage vers le haut de l'ICM-bm modélisé observé dans cette zone non ou très peu influencée par une altération trophique (PO4) **ne peut pas être la résultante d'un niveau d'évaluation plus sévère du système Français** par rapport à l'exercice d'intercalibration précédemment réalisé.

Ces TGCE du type 5 français vont jusqu'à atteindre, pour 2 d'entre eux, la valeur d'EQR-IBD de 1, ce qui correspond dans ce cas à la note maximale d'IBD₂₀₀₇ de 20 (2 relevés de la Durance, TGCE alpin venant directement de la haute montagne). Dans le même temps, ils atteignent la valeur très élevée de 1,6 dans l'échelle de notation de l'ICM-bm (pour un niveau de référence de 0,83 fixé pour cet étalon commun modélisé lors de l'exercice d'intercalibration X-GIG précédent).

Ces valeurs exceptionnellement élevées atteintes par l'ICM-bm dans des situations Françaises (à mettre en comparaison avec un seul échantillon atteignant 1,45 dans tout le jeu de données du 1^{er} exercice X-GIG) témoignent tout simplement d'une spécificité du jeu de données français par rapport aux résultats d'ensemble du 1^{er} X-GIG, illustrant notamment que ce type alpin englobe des flores d'une très bonne qualité biologique qui n'ont été qu'exceptionnellement rencontrées à cette échelle de taille de cours d'eau dans le jeu de données regroupé du précédent X-GIG.

Le calage des classes de qualité dans ce référentiel modélisé ICM-bm pourrait donc ne pas être correctement représentatif de la réalité biologique de certains des TGCE Français. En tout cas, au moins pour le groupe de sites très naturels préservés du gradient d'altération, le très haut niveau d'évaluation biologique procuré selon les 2 systèmes d'évaluation diatomique mis en relation est concordant, ce qui est plutôt rassurant.

Cette particularité de positionnement inédit des valeurs de l'ICM-bm sur des situations en bon à très bon état est donc à relier à une lacune relative, dans le jeu de données du 1^{er} X-GIG utilisé pour l'apprentissage de l'ICM-bm modélisé, en relevés TGCE représentatifs de situations d'aussi bonne qualité hydrobiologique que l'on peut en trouver en France. Or cet état de fait non lié à un contexte d'altération et d'évaluation de l'altération par le PO4 influe, au moins de ce côté des très basses valeurs du gradient de phosphore, dans le sens d'un **décalage vers le haut** de l'ajustement linéaire obtenu entre les EQR-IBD et l'ICM-bm pour le jeu de données Français.

En élément de confirmation, à une échelle plus européenne, que des communautés diatomiques rencontrées en cours d'eau de contexte montagnard sont susceptibles d'atteindre un niveau de qualité exceptionnellement élevé pour l'échelle de taille TGCE, il est possible de constater, vers le haut et la droite de la Figure 44, que quelques relevés de cours d'eau Italiens (figurés en couleur noire) collectés dans le même type de contexte que les cours d'eau français précités (cours d'eau de pied-mont ayant une origine montagnarde proche) arrivent à peu près sur le même type de positionnement.

Le fait que, contrairement à l'EQR-IBD, les valeurs d'ICM-bm (modélisé) s'envolent vers de très fortes valeurs jamais rencontrées lors de l'exercice X-GIG précédent, alors qu'il s'agit ici de contextes naturels, n'est donc pas à interpréter ici comme le signe d'un système d'évaluation Français plus sévère que d'autres systèmes nationaux, mais témoigne simplement que le calage du modèle de l'ICM-bm a été élaboré en l'absence ou quasi-absence de situations illustratives de TGCE présentant des flores d'aussi bonne qualité biologique (cf. Figure 44 : un seul point bleu de l'exercice précédent atteignant une valeur de 1,45), comme on peut en rencontrer notamment en France ou en Italie.

Ce cas de figure est naturellement le plus marquant pour des TGCE situés en sortie immédiate de contextes montagnards de haute altitude (ex : Durance), très particuliers pour cette échelle de taille de cours d'eau et de meilleure qualité biologique, en moyenne, que la plupart des TGCE rencontrés à l'échelle européenne. Cependant, outre ces 4 relevés appartenant au type 5 (Alpin proche), qui atteignent les valeurs les plus exceptionnelles d'ICM-bm, on peut aussi remarquer que tout un alignement de relevés appartenant au Type 2 (Vilaine, 1 relevé), au Type 4 (rivières carbonatées de plaine, une quinzaine de relevés) et plus exceptionnellement au Type 6 (TGCE d'origine montagnarde un peu plus éloignée, 1 relevé) subissent aussi, face à les notes d'EQR-IBD témoignant d'une qualité de bonne à très bonne des notations biologiques, mais restant toutefois dans une gamme contrôlée, des notations exceptionnellement élevées et inédites au niveau de l'ICM-bm modélisé.

A ce stade, il faut rappeler que le **niveau de référence de l'ICM-bm** a été calé à **0,83** sur la base de l'exercice d'intercalibration précédent. Les éléments d'analyse présentés ci-dessus tendent à montrer que le constat de réponse exceptionnellement élevée de l'ICM-bm pour certains relevés TGCE de France n'est tout d'abord pas «type-specific » (*i.e.* limité à un seul type particulier de TGCE Français), et que l'étalon commun benchmark (ICM-bm) n'a probablement pas été suffisamment confronté à des flores de bon à très bon niveau de qualité biologique lors de sa phase de modélisation initiale (réalisée à partir de l'assise de relevés diatomiques nationaux constitutifs du jeu de données utilisé pour le 1^{er} exercice d'intercalibration X-GIG).

Pendant toute la poursuite du processus d'intercalibration, dont la procédure est grandement basée sur des décalages d'ajustements linéaires des évaluations nationales exprimées en EQRs avec l'ICM-bm, il sera important de garder en mémoire : 1) l'ajustement plutôt de type curvilinéaire de notre propre système national avec l'étalon européen modélisé ICM-bm (cas de figure ne faisant pas l'objet de recommandation méthodologique particulière pour la procédure de Fit-in) ; 2) le décalage des échelles de notation entre les 2 outils (EQR-IBD et ICM-bm), ainsi que 3) le fait qu'une partie de ce décalage qui émane de sites non-altérés, donc hors de toute considération de mesure d'altération, entraîne des conséquences sur la globalité de l'ajustement linéaire produit (participation à la pente et au décalage de la relation entre les 2 métriques).

5.4.6. 2^{ème} itération du processus d'intercalibration (V2 du prototype d'évaluation des TGCE de France)

a) *Principes de calage du Scénario 2 et de ses 2 variantes :*

Il est tout d'abord à noter que cette 2^{ème} itération, effectuée cette fois avec le jeu de données Français définitif tel qu'envoyé le 12-12-2015, a constitué la **première occasion réelle** de confronter notre prototype de système national avec l'ICM-bm, c'est-à-dire non-entachée de l'erreur de version sur notre jeu de données national.

Ce scénario N° 2, présenté en **Figure 48** ci-dessous, est resté sur une base typologique comparable à celle du premier dispositif testé (présenté en Figure 34). En effet, il se base aussi sur 4 types nationaux de TGCE et tente, selon la logique antérieure qui avait prévalu pour tout le reste du dispositif Français d'évaluation diatomique des cours d'eau, de conserver la même grille unique de limites de classes de qualité en EQR (le niveau de référence adéquat, ainsi que la gamme maximale de variation possible par type pouvant être différents selon le type de cours d'eau et de conditions naturelles de la zone, mais le rapport des limites de classes de qualité par rapport au niveau de référence adéquat restant le même en EQR : cf. partie b de la Figure 48).

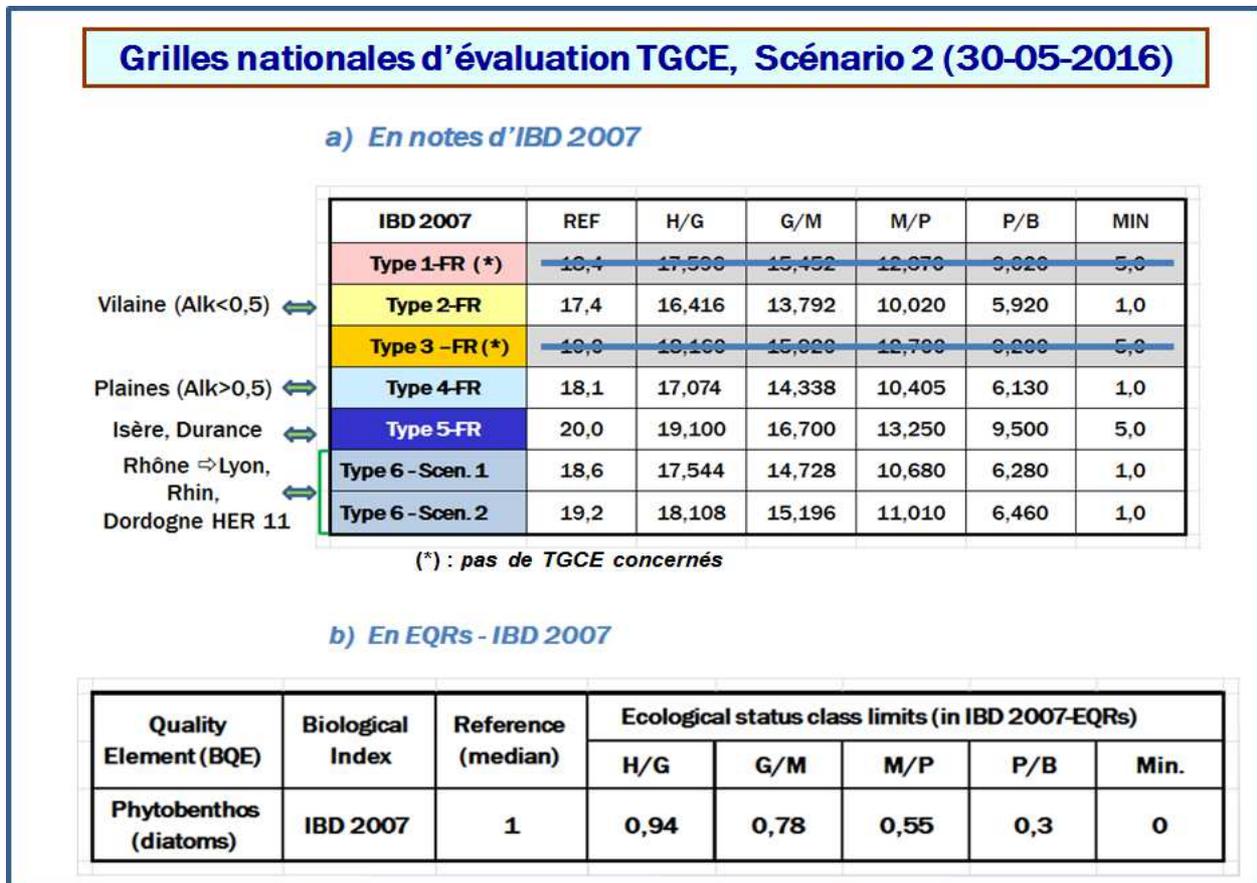


Figure 48 : Test d'intercalibration du 2ème scénario d'évaluation des TGCE de France (30/05/2016)

Les points marquants de ce scénario sont que le Type 2, le Type 4 et le Type 5 restent utilisés avec les mêmes référentiels de valeurs que dans la proposition précédente, afin d'obtenir les calculs d'ajustement et de biais exacts par rapport au jeu de données officiel Français.

Ce scénario teste l'évaluation des cours d'eau d'origine montagnarde, dont certains présentent de grosses particularités selon leur notation par l'ICM-bm, selon 2 grilles :

- la **grille Alpine (Type 5)** tel que déjà utilisée pour des cours d'eau alpins plus petits, mais dont l'utilisation est étendue vers l'aval sur de gros affluents du Rhône d'origine alpine jusqu'à leur confluence avec ce fleuve (grille utilisée sur l'Isère aval et la Durance aval).
- une **grille spécifique (Type 6)** pour des **TGCE d'origine montagnarde plus éloignée**, souvent de plus grosse taille et ayant dû traverser un gros lac naturel ou de grosses retenues à l'amont de la station (cas du Rhône, du Rhin, de la Dordogne dans l'HER 11).

Dans la mesure où, entre autres cours d'eau, les cours d'eau montagnards semblent poser un problème un peu plus exacerbé d'évaluation par l'ICM-bm, atteignant des notes témoignant d'assemblages biologiques d'un niveau de qualité inédit dans l'exercice précédent du X-GIG « large Rivers », et en vue de mieux préparer la soumission d'un dispositif national définitif, 2 variantes d'évaluation du Type 6 ont été réalisées sur la base d'un niveau de référence différent ; les buts sous-jacents étant de bien comprendre quelle évolution du niveau de référence provoque quels effets dans l'exercice collectif réalisé selon les EQR, et de préparer une éventuelle fusion de grilles d'évaluation pour les cours d'eau montagnards.

Ainsi, outre la reprise exacte de la première grille d'évaluation du Type 6 présentée en Figure 34, mais en intercalibrant le bon jeu de données national, la 2^{ème} variante testée présente un positionnement intermédiaire entre l'évaluation du Type 5 (le plus sévère) et la première grille de Type 6-Scén-1 (plus laxiste). A cet effet, le niveau de référence testé ici est intermédiaire entre ces 2 grilles (19,2) et la gamme des possibles en matière de notes d'IBD est élargie par rapport aux cours d'eau alpins plus petits (de 1 à 20 au lieu de 5 à 20 pour ce qui concerne la grille 5), correspondant mieux à des caractéristiques de cours d'eau plus importants. La comparaison des 2 alternatives de grille 6 doit permettre d'orienter le choix le plus adéquat de grille(s) montagnarde(s) pour le dispositif final, en fonction de l'interprétation des effets provoqué par ces alternatives sur le positionnement de la France dans l'exercice collectif.

b) Résultats du Scénario 2 :

Les résultats présentés en **Figure 49** sont ceux produits par le Scénario 2 - Variante 2, en confrontation avec les résultats procurés par le jeu de données du premier exercice X-GIG. Cette variante serait celle qui *a priori* permettrait d'envisager le plus facilement la fusion des 2 grilles montagnardes en une seule et de simplifier ainsi l'application du dispositif réglementaire TGCE proposé pour la France.

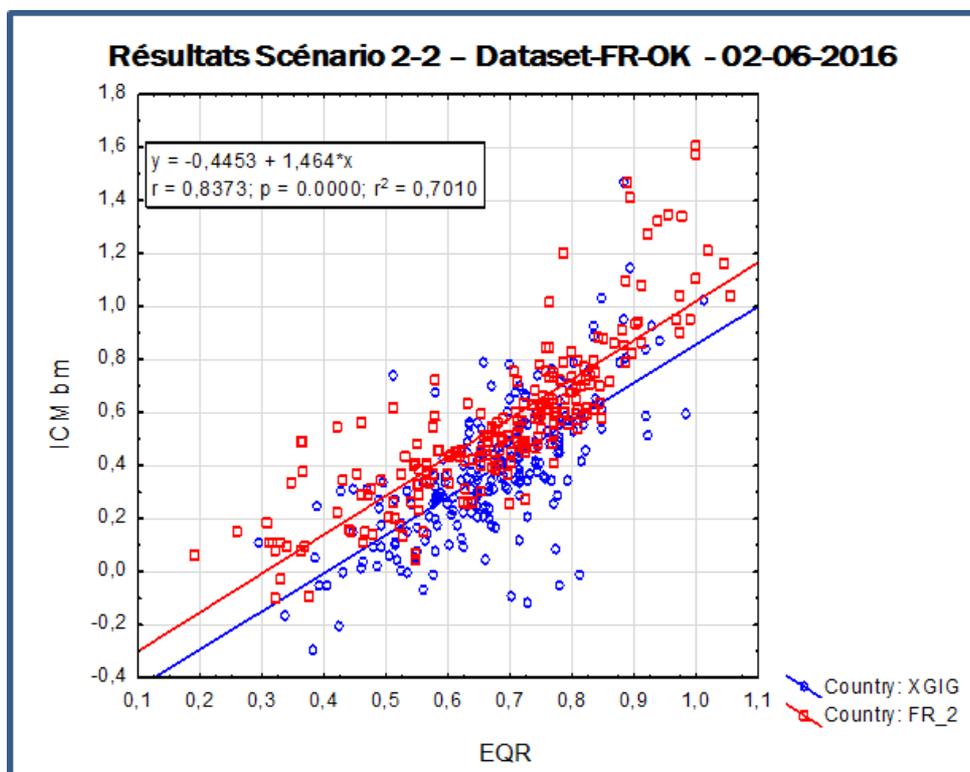


Figure 49 : Régression linéaire entre la méthode d'évaluation française (Scénario 2-2) et l'ICM-bm modélisé, en inter-comparaison avec le jeu de données du précédent X-GIG

Par rapport à la Figure 45 obtenue avec le premier jeu de données français partiellement erroné, la prise en compte du jeu de données Français validé, telle que présenté en Figure 49, n'a modifié que de façon anecdotique la physionomie globale de l'inter-étalonnage de notre dispositif national par rapport aux résultats de l'exercice X-GIG précédent, ainsi que l'équation obtenue et le calcul des résidus (R^2 quasiment identiques). La raison principale de ces faibles différences globales s'explique par l'aspect minoritaire de la proportion de relevés de cours d'eau montagnards dans le jeu de données complet fourni.

De plus, les modifications apportées à l'évaluation du Type 6 en fonction des 2 variantes Scén-1 et Scén-2 (Scén-1 non présenté car copie pratiquement conforme de la Figure 45) ne modifient qu'à la marge le positionnement des relevés montagnards en cause, du fait que les biais d'ajustement sont calculés à partir de résultats d'évaluation donnés en EQRs nationaux. Dans ce contexte, augmenter ou baisser le niveau de référence en note indicelle d'un type de TGCE donné ne va pas provoquer un effet rapidement palpable sur l'ajustement des jeux de données dans la zone des sites de référence, puisque le niveau de référence moyen (plus ou moins sévère) sera dans tous les cas de figure renormé à la valeur de 1 dans l'exercice d'inter-étalonnage.

Dans le cadre des exercices de Fit-in (intercalibration d'un nouveau système national sur la base des résultats d'un exercice précédent d'intercalibration validé), la coordinatrice de l'exercice est tenue d'appliquer scrupuleusement les consignes méthodologiques et d'interprétation données par la Guidance ECOSTAT N° 30. En fonction du cadre interprétatif donné dans ce document, le décalage vers le haut du système d'évaluation français (en rouge) serait l'expression d'un système plus sévère que la vision dominante de l'état écologique issue du 1^{er} exercice X-GIG (en bleu).

c) Biais obtenus et répercussions pour le dispositif Français :

Sur la base de l'ajustement linéaire obtenu et d'un recalage effectué sur la base des percentiles 10 et 90 des 2 métriques, la phase suivante de l'intercalibration vise à mesurer les écarts (biais) présentés par la métrique nationale par rapport au jeu de données de l'exercice précédent, en particulier aux limites de classes de qualité (en particulier, à la limite High-Good et à la limite Good-Moderate, limite basse du Bon Etat et principal point stratégique d'intercomparaison des systèmes nationaux).

Selon cette guidance, des biais négatifs sont à interpréter comme le signe que le système national est moins sévère (*i.e.* plus laxiste) que la vision collective du Bon Etat. Des biais positifs sont interprétés dans le sens d'une vision plus rigoureuse du système national par rapport à la vision collective. Dans le cas des Très Grands Cours d'Eau, les biais négatifs dépassant la valeur de 0,25 subissent une **incitation à ré-alignement**. Les biais négatifs dépassant un biais de 0,5 ne permettent pas l'acceptation de l'intercalibration par l'UE et nécessitent un **réalignement obligatoire par rapport au niveau collectif de Bon Etat moyen**. Même s'il leur est formulé une incitation à ré-aligner leur système, les **pays dont le dispositif est jugé plus sévère** sont libres, s'ils le jugent justifié, de **conserver des biais positifs**.

Les résultats de calculs de biais réalisés sur la base du 2^{ème} scénario Français (Variante 2-2) sont donnés en **Figure 50** ci-dessous. Logiquement, suite au décalage observé en Figure 49, ce dispositif présenté par la France subit des biais positifs substantiels (0,64 pour la limite H-G et 0,61 pour la limite G-M). Ces 2 limites de classes dépassant la valeur de biais maximum acceptable de 0,5, ce prototype de dispositif Français s'est vu adresser une recommandation d'ajustement à la baisse.

National specificities			National boundaries EQR					Boundaries on ICM + offset			BIAS		Result		
Country	Offsets	*national type	H/G	G/M	M/P	REF	P/B	H/G on ICM	G/M on ICM	M/P on ICM	REF	H/G bias	G/M bias	H/G	G/M
FR_Scen 2-2	0,177092	IBD-EQR (every type)	0,94	0,78	0,55	1	0,3	0,75376846	0,51952846	0,18280846	0,8416	0,643649	0,616324	passed*	passed*

* adjustment strongly recommended

Figure 50 : Biais d'intercalibration calculés pour le Scénario 2-2 de dispositif Français d'évaluation TGCE

Même s'il apparaît que certaines évaluations Françaises sont assez sensiblement décalées vers le haut, dans l'échelle de l'ICM-bm, par rapport à l'apprentissage initial réalisé sur la base de l'exercice X-GIG précédent, il convient de rester prudent sur les raisons susceptibles d'occasionner et d'interpréter ces décalages, qui plus est sur la base d'un cadre interprétatif défini *a priori* et qui s'appuie sur une forte assomption de linéarité des relations obtenues, mais qui n'avait pas envisagé (et peut donc difficilement répondre de façon satisfaisante à) d'autres types d'ajustement possibles.

D'autre part, dans le cadre du présent exercice de Fit-in, il a pu être nettement constaté que l'ICM-bm pouvait, en fonction de son jeu de données d'apprentissage, donner **des évaluations nettement décalées, y compris sur des flores non-altérées** et de ce fait, contribuer sensiblement à un décalage de l'ajustement linéaire obtenus entre une métrique nationale et le jeu de données complet de l'exercice précédent sans que, sur cette partie de l'ajustement, une évaluation trop sévère ne puisse être la raison du décalage observé.

Même s'il faut éviter que la France n'évalue les TGCE d'une façon systématiquement plus sévère que la vision dominante ressortant des autres pays-Membres, ce qui contribuerait à dégrader l'état écologique évalué et à alourdir le poids des programmes de remédiation, l'interprétation de ces résultats bruts du 2^{ème} scénario se confronte aux éléments modérateurs déjà évoqués précédemment :

- l'IBD2007, qui sert de base à notre système d'évaluation national, n'évalue pas que l'altération unique par le PO4 considérée dans cet exercice, mais intègre de façon composite les effets occasionnés par 7 descripteurs différents de pression anthropique sur la qualité de l'eau ;
- .en fonction d'une réponse très haute et inédite de l'ICM-bm, le positionnement de points très décalés en haut et à droite contribue, au moins de ce côté du gradient, à modifier la pente de l'ajustement et à décaler vers le haut l'ajustement Français. Or, sur cette partie du graphe hors anthropisation, ces très hautes valeurs d'ICM-bm n'ont aucun rapport avec une sévérité excessive d'évaluation de l'IBD2007.

- Au sein des Pays-Membres, la façon d’appréhender le niveau de référence par défaut et de construire les grilles d’EQR a été manifestement hétérogène, certaines grilles nationales manifestement bien calées conduisant à des valeurs maximales d’EQR pouvant atteindre ou dépasser modérément la valeur de 1, mais d’autres systèmes plus laxistes, qui sous-estiment nettement la situation de référence pour le type, pouvant conduire à certains calculs d’EQR dépassant 1,4 ou 1,5 sur des sites à évaluer. Or, dans l’exercice d’intercalibration, ce sont les valeurs d’évaluation de systèmes nationaux **données en EQR** qui servent de base aux ajustements linéaires et aux calculs de biais, ce qui est susceptible de biaiser fortement les résultats obtenus par un pays ou par un autre.

A ces réserves près, qui dépassent notre sphère directe d’action dans le cadre de l’exercice, mais incitent à une certaine prudence par rapport à la révélation de décalages et aux raisons multiples susceptibles d’y contribuer, la France a tout intérêt à évaluer ses hydrosystèmes de façon équilibrée par rapport aux autres Pays-Membres, donc à revenir à un niveau d’évaluation plus proche de la vision dominante ressortant de l’exercice.

Le 3ème Scénario a donc été réfléchi dans la double intention de simplifier le dispositif réglementaire à appliquer et de diminuer un peu la sévérité du prototype national d’évaluation, en vue de converger vers une vision plus partagée du bon Etat.

5.4.7. 3^{ème} itération du processus d’intercalibration (V3 du prototype d’évaluation des TGCE de France)

Cette 3ème version de prototype, qui a fait l’objet du test d’un tunage fin en vue d’affiner le calage de notre dispositif national définitif, a été élaborée et testée entre le 3 Juin et le 16 Juin 2016, soit moins d’une semaine avant l’envoi par la coordinatrice du rapport final d’intercalibration « Fit-in exercice » à l’UE (21-06-2016).

a) *Analyse des problèmes rencontrés, solutions envisagées*

1) **Cas des cours d’eau alpins et montagnards :**

Dans l’état, certains cours d’eau français ont un niveau de qualité biologique jugé extrêmement élevé par l’ICM-bm (cf. secteur en haut à droite de la Figure 47), dont en particulier (mais pas uniquement) des cours d’eau du Type 5 (origine Alpine proche).

Le décalage vers le haut qui en résulte se rencontre sur certains sites parfaitement naturels, où ce décalage vers de très fortes valeurs n’est donc pas explicable par une plus grande sévérité d’évaluation de notre dispositif national, mais dans un apprentissage de l’ICM-bm effectué avec un jeu de données présentant un défaut de situations d’aussi bonne qualité hydrobiologique (donc un défaut de représentativité par rapport à certaines situations naturelles effectivement rencontrées dans le présent exercice de Fit-in). Cependant, sur des sites impactés à évaluer, il pourrait aussi conduire à des évaluations d’altérations plus sévères en EQR-IBD du fait que cette grille d’évaluation de Type 5 (cours d’eau Alpains) est raccourcie vers le bas. En effet (cf. Figure 48), l’EQR 0 pour le Type 5 correspond à une note minimale d’IBD de 5, valeur minimale atteignable par cet indice dans des cours d’eau du type torrents de montagne à forte hydrodynamique.

Dans le cas présent, comme les TGCE montagnards sont par définition de taille plus importante et s’écoulent à des altitudes un peu plus basses que les torrents d’échelle de taille inférieure, il serait envisageable et logique de diminuer un peu le niveau de référence (ici, par défaut) de tels cours d’eau en notes d’IBD₂₀₀₇, ce qui devrait contribuer à réduire un peu la sévérité du dispositif d’évaluation. D’autre part, le fait de les évaluer, comme la plupart des autres cours d’eau de France, sur une gamme plus étendue de notes possibles d’IBD (l’EQR 0 correspondant dans ce cas à une note minimale d’IBD de 1), serait de nature à diminuer la sévérité d’évaluation en situations altérées et à se rapprocher de la vision dominante du Bon Etat pour les TGCE. Il a donc été testé l’impact d’une simplification de grilles de TGCE alpins et montagnards, qui pourrait conduire à l’évaluation de tous les TGCE d’origine montagnarde proche sur la base de la seule grille 6 (abandon de la grille du Type 5).

La partie a), à gauche de la **Figure 51** (page suivante) présente le calage de l’évaluation des TGCE Alpains selon la grille de type 5 utilisée dans le cadre du Scénario 2 testé précédemment (Figure 48, Type 5-FR). Il est visible que, par rapport à l’ajustement entre EQR-IBD et ICM-bm obtenu pour l’ensemble du jeu de données Français sur la base du jeu de données complet du 1^{er} exercice X-GIG (ajustement linéaire figuré en rouge), cette grille de type Alpin produit une évaluation nettement décalée vers le haut, contribuant à faire évaluer notre système français comme trop sévère selon le cadre interprétatif de la Guidance Européenne. Pour le même ensemble de relevés, le fait de passer sur une grille d’évaluation de type 6 (**Figure 51**, partie b à droite) conduit à une translation horizontale vers la droite des résultats de cette intercomparaison (les points rouges glissent vers la droite et deviennent les points bleus), se rapprochant ainsi de l’ajustement linéaire général et conduisant à un recentrage des résidus de ce groupe de relevés.

Ce constat prévaut sur le plan général, à l'exception de 4 relevés d'extrêmement bonne qualité écologique qui correspondent à des situations atypiques par rapport à la phase d'apprentissage de l'ICM-bm.

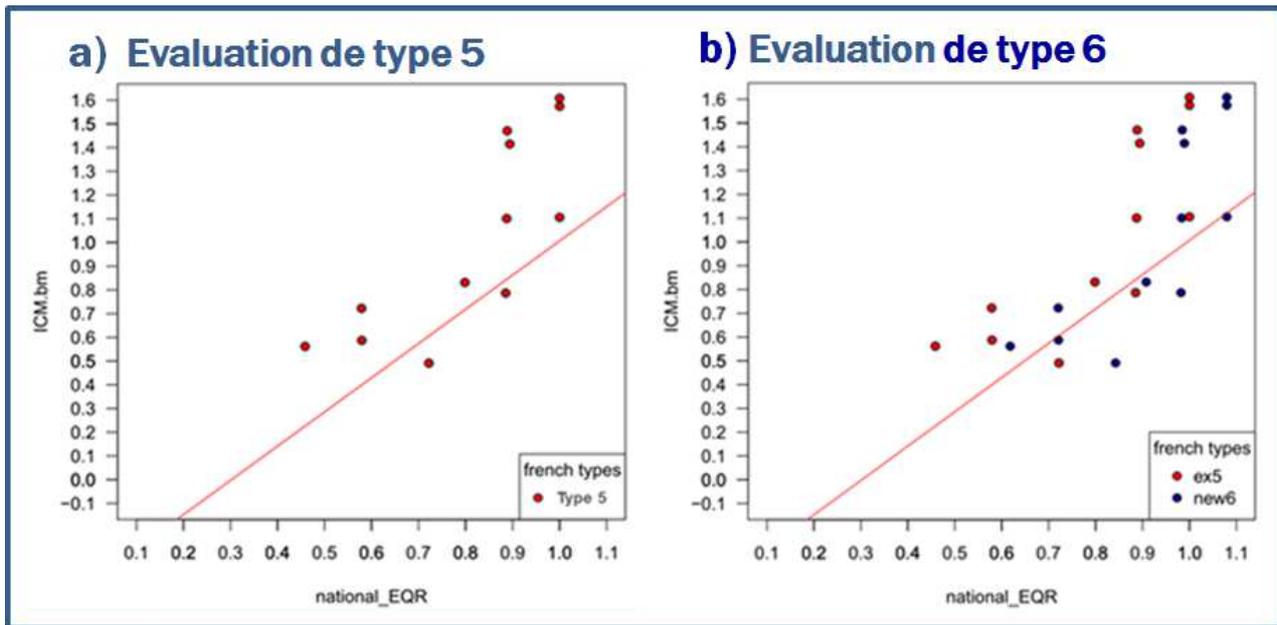


Figure 51 : Impact d'un changement de grille d'évaluation pour les TGCE du Type 5 (Alpin)

Une telle évolution de dispositif serait donc favorable à une évaluation plus DCE-compatible de ce groupe de cours d'eau alpins de grande taille, mais pouvant présenter des caractéristiques torrentielles (les relevés en cause sont ceux de l'Isère et de la Durance).

Cependant, vu que seulement 12 relevés du jeu de données national d'intercalibration sont concernés par ce changement de type (par rapport à un total de 221 relevés diatomiques, soit seulement 5%), la modification globale d'ajustement d'évaluation qui résulterait de ce seul changement de grille restera forcément marginale et ne suffira probablement pas pour ramener les biais de notre dispositif national dans les limites jugées acceptables pour l'exercice (les biais jugés tolérables devant être de moins de 0,5).

2) Réflexion en vue d'un réalignement plus général :

Jusqu'à présent, notre dispositif national d'évaluation diatomique des cours d'eau a été construit selon le principe général que l'évaluation devait tenir compte des particularités naturelles des zones homogènes concernées, à savoir un niveau de référence moyen calculé à partir des notes d'indice obtenues en situations de référence de la zone concernée (qui équivaut à l'EQR de 1) et une note indicelle minimale atteignable en cas de très forte altération dans la zone naturelle considérée (qui équivaut à l'EQR de 0). Dans ce cadre naturel variable, une seule et même **grille d'évaluation en EQR était utilisée en France**, traduisant un écart standardisé par rapport au niveau de référence calculé pour la zone naturelle concernée.

Dans la mesure où l'Union Européenne n'a pas donné de cadrage initial très précis aux Etats-Membres, dont certains disposaient déjà de dispositifs hydrobiologiques d'évaluation pré-existants avant l'entrée réelle en vigueur de la DCE (par exemple, alors que la France avait déjà des réseaux de surveillance diatomique depuis les années 1995-96, la 1^{ère} circulaire nationale d'évaluation découlant de la mise en œuvre de la DCE y est entrée en vigueur en Juillet 2005), la logique de construction des EQR adoptée par les différents Pays-Membres n'a pas forcément été comparable, conduisant à des principes d'évaluation hétérogènes et à des anomalies qui, si elles ne font pas l'objet d'une attention spécifique, peuvent impacter les exercices d'inter-comparaison. Par exemple, des EQrs nationaux peuvent atteindre plus de 1,4 ou 1,5 pour certains pays, alors que d'autres restent cantonnés à un EQR maximum atteignable très peu supérieur à 1 (la France se classe dans ce dernier cas de figure), ce qui entraîne inévitablement de grosses distorsions d'échelle et de niveau d'évaluation.

Dans le cas où un site à évaluer obtient un EQR de 1 ou très proche de 1, au sens de la DCE, le site ne révèle pas d'écart par rapport au niveau de référence adéquat pour sa zone naturelle et sera donc jugé en très bon état. Mais, sur le plan écologique et de mesure de l'altération écologique, que conclure d'un site à évaluer qui obtient une valeur de 1,5 fois le niveau de référence moyen pour la zone, si ce n'est que le système national est probablement trop laxiste ou, du moins, que la zone naturelle d'application d'une même grille en EQR n'est pas suffisamment homogène sur le plan de ses variabilités naturelles internes.

Dans l'état, les évaluations nationales faites avec des systèmes basés sur des philosophies de construction si différentes s'effectuent donc sur une base difficilement comparable, et de telles différences sont susceptible d'impacter sensiblement la validité de l'intercomparaison réalisée entre systèmes nationaux...

Dans la logique de l'intercalibration européenne, le seul facteur d'atténuation de ce phénomène est qu'un pays à pratique d'évaluation marginale, et éventuellement à positionnement marginal, même s'il contribue à tirer un peu vers lui la situation moyenne du jeu de données collectif, exercera un effet limité sur la fixation de la vision commune du Bon Etat, l'intercalibration étant réalisée par rapport à la vision dominante de tout un groupe (dans le cas de notre exercice de Fit-in, ce niveau se cale par rapport à la vision dominante résultant de tout l'exercice précédent, soit le 1^{er} exercice du X-GIG Large Rivers).

Dans le cas présent, la réalisation de l'exercice d'intercomparaison s'effectue sur la base des **valeurs en EQR nationaux** dérivées des notes indicielles brutes. Même si, dans l'absolu, une pratique de diminution (en note indicielle) du niveau de référence visé pour une zone naturelle donnée doit conduire à une moindre sévérité du dispositif d'évaluation concerné, il ne serait pas efficient d'essayer d'ajuster des niveaux de référence par défaut (exprimés en note moyenne d'IBD₂₀₀₇) à des valeurs plus basses par type de TGCE. En effet, le niveau de référence réduit en note indicielle servirait de base à la renormalisation en EQR et la valeur de référence moyenne exprimée en EQR s'établirait à nouveau à la valeur de 1. Au-delà du possible changement de pente de l'ajustement selon les 2 métriques (dont on aurait éventuellement du mal à interpréter la signification en termes de sévérité ou de laxisme d'évaluation, vu l'aspect non linéaire de la relation entre IBD₂₀₀₇ et ICM-bm), les biais d'évaluation ne s'en trouveraient que très modérément abaissés à la limite de chaque classe de qualité.

En fonction de la méthodologie d'intercalibration appliquée, qui se base sur les valeurs d'EQR nationaux obtenues, la seule **solution efficace** pour diminuer les biais positifs actuellement calculés a donc consisté à **proposer une échelle d'évaluation en EQR différente, pour les TGCE**, que pour toutes les autres catégories de cours d'eau plus petites (déjà intercalibrées OK préalablement, donc basées sur une grille ayant force d'application), à savoir de proposer des limites de classes de qualité en EQR rabaissées par rapport au niveau de référence (1) et par rapport à la grille d'EQR unique utilisée pour les cours d'eau plus petits.

Au plan national, ce cas de figure contraint par des aspects méthodologiques de l'intercalibration n'était pas souhaité, car le principe de construction et d'utilisation d'EQRs ne reste pas homogène avec celui qui a prévalu pour tous les autres cours d'eau et introduit une nouvelle complexité dans le dispositif (2 grilles différentes d'évaluation en EQRs au lieu d'une seule auparavant).

b) 3^{ème} Scénario de dispositif national proposé (03-06-2016)

Pour la construction de ce 3^{ème} Scénario (voir **Figure 52**), les TGCE alpins (Type 5) ont été fusionnés dans le même ensemble typologique que les cours d'eau montagnards du Type 6. Il a aussi été procédé à une diminution marginale du niveau de référence pour ce nouveau Type 6 « regroupé » (soit 19,1 au lieu de 19,2 dans le Scénario 2). Enfin, en vue de se rapprocher plus sensiblement de la vision partagée du Bon Etat obtenue dans le 1^{er} exercice d'intercalibration du X-GIG « Large Rivers », ce nouveau prototype de dispositif national utilise une **nouvelle échelle d'évaluation en EQR dédiée spécifiquement à tous les types de TGCE**, qui présente des limites de classes de qualité un peu abaissées par rapport à la grille unique utilisée au niveau national pour les cours d'eau plus petits (cf. grille d'évaluation en EQR en bas de la Figure 52, au lieu de celle du bas de la Figure 30).

Grille d'évaluation proposée (Scénario 3-FR, 03-06-2016)

IBD 2007	REF	H/G	G/M	M/P	P/B	MIN
Region 2 - Scen. 3	17,4	16,088	13,546	9,528	5,264	1,0
Region 4 - Scen. 3	18,1	16,732	14,082	9,892	5,446	1,0
Region 6 - Scen. 3	19,1	17,652	14,847	10,412	5,706	1,0
EQR-IBD	1	0,92	0,765	0,52	0,26	0

Figure 52 : Grille d'évaluation des TGCE de France selon le Scénario 3 (proposition du 03-06-2016)

Ce nouveau dispositif proposé a fait l'objet d'un retour rapide de calcul de biais par la coordinatrice, permettant de situer les progrès effectués en matière de niveau d'évaluation en vue de la proposition Française finale.

Boundary bias comparison (Scenario 3-FR / Scenario 1-FR)															
National specificities			National boundaries EQR					Boundaries on ICM + offset				BIAS		Result	
Country - Scenario	Offsets	national type	H/G	G/M	M/P	REF	P/B	H/G on ICM	G/M on ICM	M/P on ICM	REF	H/G bias	G/M bias	H/G	G/M
FR - Scenario 1	0.2047	IBD-EQR (every type)	0.94	0.78	0.55	1	0.5	0.715355	0.484555	0.15278	0.8	0.4868	0.5203	passed	passed
FR - Scenario 3	0.2002	IBD-EQR (every type)	0.92	0.765	0.52	1	0.26	0.691698	0.464206	0.104698	0.81	0.3899	0.4234	passed	passed

* adjustment strongly recommended

National OLS regression (y=ICM, x=EQR):
Scenario 1 $y = 1.44254x - 0.4359$
Scenario 3 $y = 1.4675x - 0.4552$

Figure 53 : Dispositif Français, Scénario 3 : Evolution des biais par rapport au Scénario 1

Il a ainsi été possible de constater via l'évolution des régressions OLS une pente d'ajustement linéaire en très légère augmentation entre les 2 métriques. L'autre évolution positive est que les biais calculés pour les 2 limites de classes H/G et G/M (ligne la plus basse de la Figure 53) ont assez sensiblement diminué par rapport au Scénario 1. Tout en restant assez nettement positifs (ce qui, malgré plusieurs réserves méthodologiques déjà discutées, est interprété dans le cadre de l'exercice comme l'expression d'une évaluation Française plus sévère que la moyenne de l'exercice de référence), ils rentrent cette fois à l'intérieur de la limite appliquée pour l'exercice (<0,5) pour ces 2 limites de classes.

Ce retour de la coordinatrice ayant été très rapide, il a été décidé avec l'AFB et la Direction de l'Eau de lui soumettre en retour rapide 3 variantes de grilles d'évaluation en EQR permettant de tester un tunage fin de ce Scénario 3, en vue d'effectuer le choix final et de pouvoir lui envoyer, avant la date limite du 20-06-2016, une proposition stabilisée de dispositif Français officiel soumis à l'exercice final d'intercalibration.

c) Tunage fin du 3^{ème} Scénario de dispositif Français d'évaluation (06 au 13-06-2016)

Du fait que l'exercice européen d'intercalibration est effectué sur la base des évaluations produites en EQR nationaux, en vue de choisir la grille unique à appliquer au final sur les TGCE nationaux, tous types confondus, 3 variantes de grilles exprimées en EQR ont été testées (voir **Figure 54** ci-dessous). Le tunage fin a seulement porté sur le calage de la limite inférieure du Bon Etat (limite Good / Moderate), selon la gradation de seuils proposée dans l'encadré rouge ; les nouvelles limites testées venant encadrer le Scénario 3-2 (limite G/M fixée à 0,765), qui correspond exactement au Scénario 3 proposé à l'étape précédente.

Tunage fin des grilles d'EQR (Scénario 3-FR, 06 au 13-06-2016)

Tunage fin EQR-IBD	REF	H/G	G/M	M/P	P/B	MIN
Scénario 3-1	1	0,92	0,77	0,52	0,26	0
Scénario 3-2	1	0,92	0,765	0,52	0,26	0
Scénario 3-3	1	0,92	0,76	0,52	0,26	0

Figure 54 : Essai de tunage fin de grilles d'EQR-IBD2007 dérivées du scénario 3 (06 au 13-06-2016)

En date du 13-06-2016, la Coordinatrice nous a renvoyé ses calculs de biais de ces 3 variantes de Scénario 3 par rapport à l'ajustement du jeu de données X-GIG précédent (voir **Figure 55** en page suivante).

Boundary bias comparison between Scenario 3-FR and Scenario 1-FR (13-06-2016)															
National specificities			National boundaries EQR					Boundaries on ICM + offset				BIAS		Result	
Country - Scenario	Offsets	*national type	H/G	G/M	M/P	REF	P/B	REF	H/G on ICM	G/M on ICM	M/P on ICM	H/G bias	G/M bias	H/G	G/M
FR - Scenario 1	0.2047	IBD-EQR (every type)	0.94	0.78	0.55	1	0.5	0.8019	0.7154	0.4846	0.1528	0.4868	0.5203	passed	passed
FR3 - Scenario 3-1	0.2052	IBD-EQR (every type)	0.92	0.77	0.52	1	0.26	0.8091	0.6917	0.4716	0.1047	0.40295	0.43495	passed	passed
FR3 - Scenario 3-2	0.2052	IBD-EQR (every type)	0.92	0.765	0.52	1	0.26	0.8091	0.6917	0.4642	0.1047	0.38995	0.42342	passed	passed
FR3 - Scenario 3-3	0.2052	IBD-EQR (every type)	0.92	0.76	0.52	1	0.26	0.8091	0.6917	0.4569	0.1047	0.37776	0.41141	passed	passed

* adjustment strongly recommended

Figure 55 : Calcul de biais sur les 3 variantes du Scénario 3-FR (13-06-2016)

Pour mémoire, le scénario initial rappelé en première ligne (Scénario 1) se situait au-delà du biais maximum de 0,5 jugé acceptable pour l'exercice, et les biais positifs calculés étaient interprétés comme l'expression d'un dispositif national plus sévère que la vision dominante du Bon Etat. Comme attendu, les 3 variantes de Scénario 3 proposées, basées sur une baisse des valeurs de limites de classes de qualité exprimées en EQR nationaux, ont permis de diminuer les biais et de les faire rentrer assez largement dans les biais jugés acceptables pour l'exercice ; le scénario fixant la limite G/M la plus basse (Scénario 3-3, limite G/M à 0,76) étant bien sûr celui qui permet de se rapprocher le plus de la vision dominante du Bon Etat établie lors de l'exercice précédent.

La date d'envoi de ces nouveaux calculs de biais nous ayant été annoncée quelques jours à l'avance, nous avons pu tenir dès réception du résultat des tests, soit le 13-06-2016, une visioconférence pour discussion avec le représentant de la Direction de l'Eau (J.P. Cabaret) et celui de l'AFB (Y. Reyjol) afin d'examiner ces résultats et de procéder au calage final du dispositif Français officiellement présenté à l'intercalibration.

d) Calage final du dispositif Français d'évaluation des TGCE (du 13 au 16-06-2016):

Avant que la décision du niveau final de calage de notre système national soit prise pour la réalisation de l'exercice consolidé d'intercalibration, il a été refait le point du déroulement pas à pas des procédures déroulées par la coordinatrice afin d'exécuter l'exercice collectif. Les procédures méthodologiques sont pré-définies et très codifiées (Guidance 30) en fonction du cas de figure général rencontré pour la réalisation d'un exercice donné d'intercalibration. Le but étant de pouvoir bénéficier au final du visa OK de l'Union Européenne, leur application ne laisse pas forcément beaucoup de place à l'adaptation alors que des cas particuliers, qui n'ont pas forcément été envisagés *a priori*, peuvent se présenter. Dans ce cadre contraint, certaines **limites méthodologiques** concernant les règles de mise en œuvre du présent exercice ont été ré-évoquées, à savoir :

- Le **gradient unique d'altération anthropique** pris en compte dans cet exercice d'intercalibration UE a été le **PO4**. Or, si certaines méthodes nationales, effectivement, sont fléchées pour évaluer plus spécifiquement cette altération (par exemple le TDI au Royaume Uni, la métrique RTI en Autriche, qui entre d'autre part pour moitié dans l'élaboration de l'étalon européen utilisé ici, l'ICM-bm diatomique...), il en va différemment pour d'autres systèmes nationaux d'évaluation. Dans le cas de la France, par construction, l'IBD₂₀₀₇ évalue les altérations liées à **7 descripteurs différents de pression anthropique**. Sur son territoire, beaucoup de sites des réseaux de surveillance, notamment situés à l'aval d'implantations successives d'agglomérations au long des cours d'eau et/ou de STEP, sont porteurs de **gradients composites d'altération anthropique** faisant intervenir en même temps une charge organique et carbonée, plusieurs types de nutriments, de la demande en O₂, etc... Il est donc normal que notre indice national puisse donner un résultat d'altération biologique plus important que celui qui résulte de la seule présence d'orthophosphates au site concerné.
- Sur la base des éléments concrets dont nous avons pu disposer, il n'est pas complètement possible d'assurer que les **différences de niveau de gradient de PO4 observées entre les 2 exercices** (1^{er} exercice du X-GIG et présent exercice de Fit-in), qui ont fait l'objet à plusieurs reprises de remarques et de mise en garde via l'envoi de notes nationales de réactions techniques et d'expertise, aient été effectivement réglées de façon satisfaisante et définitive. Ces différences importantes de pentes d'ajustement entre l'ensemble du précédent X-GIG et les jeux de données nationaux du présent exercice (cf. Figure 44), alors qu'une partie des grands TGCE Européens représentés dans les 2 exercices sont les mêmes, ne semblent pas forcément évidentes ni justifiées à l'observation des données élémentaires d'illustration des gradients concernés pour les 2 exercices d'intercalibration (cf. partie gauche de la Figure 42). Il reste donc possible que des différences de prise en compte des référentiels de valeurs (P-PO4 versus PO4) et/ou de transformations log appliquées puissent participer aux nettes différences de pentes observées, de manière générale, entre les 2 exercices (Figure 44).

- Si les différences de pente observées entre les 2 exercices ont une part d'origine artificielle (cf. notamment les réponses anormalement élevées de l'ICM-bm sur des sites de très bonne qualité hydrobiologique, qui n'étaient pas rencontrés dans le jeu de données d'apprentissage), la solution utilisée pour recalibrer les 2 exercices, même si elle est décrite dans une Guidance Européenne de cadrage de «Fit-in exercise», peut être **entachée de biais susceptibles de modifier l'ajustement des pentes** entre les 2 exercices, donc les ajustements finaux obtenus entre les percentiles 10 et 90 des 2 systèmes intercomparés.
- L'intercomparaison avec l'exercice précédent X-GIG a clairement montré que la **réponse de l'étalon commun d'intercalibration modélisé** (ICM-bm), sur la base de l'ensemble des données nationales collectées pour l'exercice précédent (jeu de données X-GIG), **n'a jamais été confrontée à des situations de qualité biologique de TGCE aussi bonnes** que ce qui peut être rencontré dans certaines situations de France (Durance, Dordogne) ou d'Italie directement issues de zones montagnardes. Le niveau de qualité hydrobiologique de ces cours d'eau est nettement plus élevé que celui de fleuves de très grande taille à l'échelle européenne et dont de larges secteurs d'écoulements se situent en zones de plaines, comme par exemple le Danube ou la partie aval du Rhin en Allemagne puis aux Pays-Bas, entre autres exemples. La réponse de cet ICM a donc du mal à être correctement calée dans ce nouveau domaine ayant visiblement participé de façon très insuffisante (1 seul relevé) à son apprentissage et à son calage initial. Dans ces situations, c'est bien l'étalon européen (et non les EQR de méthodes nationales) qui donne des valeurs extrêmement élevées sortant de sa gamme habituelle d'interprétation, ce qui ne peut pas, au moins de ce côté du gradient correspondant aux situations naturelles, être interprété comme le signe d'une méthode nationale trop sévère (même en EQR de métriques nationales, les notes y atteignent le maximum de l'échelle possible, soit 1 ou des valeurs modérément supérieures à 1, niveau moyen de référence).
- La méthode d'intercomparaison interprète les **décalages d'ajustements linéaires** entre les EQR de métrique nationale et l'ICM-bm comme une **mesure de comparaison de sévérité ou de laxisme d'évaluation** par rapport à la vision dominante du Bon Etat Ecologique. Par rapport à l'application de ce principe, une réserve peut être émise sur les différences d'objectif de qualité établies, pour chacun des systèmes nationaux d'évaluation, en situation de référence ou de référence par défaut, donc de pente d'ajustement non forcément liée de façon équivalente au gradient de PO4 dans le jeu de données du pays considéré. Les différences nettes de méthodologie qui semblent avoir présidé à la fixation du niveau de référence par type et à la construction des grilles d'EQRs par pays conduisent certains d'entre eux à obtenir des valeurs d'évaluation incompréhensiblement trop élevées d'EQRs nationaux, qui atteignent jusqu'à plus d'1,4 à 1,5 fois le niveau de référence, ce qui ne revêt plus aucun sens intelligible sur le plan écologique. Cet état de fait est probablement le signe d'un niveau de référence établi de façon trop laxiste, ou d'une trop importante différence typologique entre les cours d'eau évalués dans un type national donné, qui ne constituerait donc pas un ensemble suffisamment homogène pour être évalué selon une même grille...). A cette importante réserve près, cette interprétation de la signification des décalages serait justifiée à l'intérieur d'un exercice d'intercalibration donné.
- Cependant, dans notre cas, certains jeux de données nationaux du nouvel exercice à réaliser **sortent visiblement du domaine biogéographique couvert par le 1^{er} exercice**, qui ne couvrait pas d'aussi bonnes situations écologiques pour des cours d'eau de taille supérieure à 10 000 km². Ce constat est en particulier confirmé pour les cours d'eau directement issus de zones montagnardes qui ne semblent pas avoir été suffisamment représentés, à l'exception d'un seul relevé, dans l'assise du jeu de données regroupé du 1^{er} X-GIG. Nous sommes donc dans le cas où des données nationales à intercaler se situent dans une zone non-couverte et visiblement non-valable de l'apprentissage antérieur réalisé par l'étalon commun modélisé ICM-bm). Les valeurs très fortes et inusitées que cet étalon obtient génèrent, au moins du côté des situations naturelles, un décalage de l'ajustement obtenu qui va forcément entraîner un décalage de pente de l'ajustement tout entier.
- Enfin, à différentes étapes de sa mise en œuvre, le principe d'intercomparaison appliqué se base sur un principe général d'**ajustement linéaire** des relations entre EQR nationaux et l'étalon commun d'intercalibration modélisé (l'ICM-BM). Or, du moins dans le cas de la France, le patron d'ajustement entre ces 2 métriques est assez nettement de type curvilinéaire (nuage rouge de la Figure 45), un signe indiscutable de cette caractéristique étant une répartition des résidus nettement déséquilibrée aux 2 extrémités de leur distribution (Figure 46). Le cadre méthodologique à peu près vérifié pour le 2^{ème} exercice de Fit-in dans sa globalité est donc beaucoup moins adéquat dans le cas particulier du système d'évaluation français, sans que les procédures définies *a priori* ne prévoient d'autre procédure plus adaptée dans un tel cas.

En conclusion, la France est venue participer au nouvel exercice d'intercalibration TGCE en tant que **simple Pays-Membre** participant, afin d'y faire intercaler sa propre proposition de dispositif national, et **pas dans l'objectif de souligner et de régler tous les problèmes méthodologiques rencontrés** par rapport à l'application d'une Guidance Européenne de Fit-in établie par un groupe de scientifiques dûment mandatés pour son élaboration (Groupe ECOSTAT), mais pas forcément adaptée à tous les cas de figures possibles.

Sur le plan opérationnel, l'aspect primordial était que la **France puisse voir son système national validé OK dans le cadre du présent exercice**, afin de pouvoir ensuite mettre en œuvre l'évaluation de la qualité écologique de ses TGCE dans des conditions officielles validées par l'U.E. ; ceci, même si plusieurs aspects méthodologiques utilisés pour la réalisation de l'exercice restent porteurs de questionnements et de réserves, notamment par rapport à l'adéquation de cette démarche et à la parfaite validité des résultats obtenus vis-à-vis de son cas de figure particulier (cf. présence importante de relevés illustratifs d'une qualité biologique inédite dans l'exercice précédent, cf. non-linéarité de l'ajustement produit pour notre jeu de données national etc...).

Ces éléments contextuels ayant été rediscutés, l'attitude pragmatique à ce stade était de s'adapter au mieux aux résultats obtenus dans le cadre de l'exercice en cours, en vue de stabiliser un dispositif national permettant une évaluation judicieuse et équilibrée de nos TGCE, tenant compte des spécificités de leur contexte naturel.

En éléments modérateurs par rapport aux niveaux de biais obtenus, le décalage vers le haut d'un bon nombre de relevés nationaux appartenant notamment à un type montagnard, mais pas seulement, dans l'échelle de l'ICM-bm, n'est donc pas à interpréter uniquement comme l'expression d'un système national « trop » sévère, mais aussi comme la conséquence de la présence, dans le cas de nos TGCE nationaux, de cortèges d'espèces inédits, de meilleure qualité écologique que dans l'assise de cours d'eau ayant servi au premier exercice d'intercalibration X-GIG. En effet, l'ajustement est aussi et surtout tiré vers le haut dans le cadre de situations parfaitement naturelles ou pseudo-naturelles, la raison du décalage ne pouvant pas être interprétée, au moins dans cette zone, comme un niveau d'évaluation trop sévère. Un autre aspect modérateur recevable pouvant être avancé par rapport au décalage positif obtenu est que l'IBD évalue des altérations anthropiques plus diversifiées que le seul gradient de PO4, dans un contexte national où plusieurs altérations co-existent sur la plupart de nos sites de réseaux. Le niveau de sa réponse biologique n'est donc pas à caler et à interpréter uniquement en fonction du gradient de PO4 présent à chaque site.

En considérant cette fois ces résultats selon un autre angle d'analyse, même si les Etats-Membres ne sont pas obligés d'ajuster à la baisse un dispositif national générant des biais positifs (alors que le contraire est obligatoire), il ne serait pas dans l'intérêt national d'évaluer de façon excessivement sévère ses propres cours d'eau par rapport à la pratique des autres Pays-Membres, le risque étant de caler trop haut la limite inférieure du Bon Etat et de montrer ainsi, dans le cadre du rapportage Européen, une situation plus dégradée qu'elle ne l'est en réalité par rapport aux pays voisins. Une conséquence opérationnelle qui en découlerait serait d'augmenter optiquement l'effort de réhabilitation nécessaire.

La solution jugée raisonnable par les 3 partenaires associés à la décision (Direction de l'Eau, AFB, Irstea) a donc été, dans le cadre du tunage fin testé, de choisir la variante de dispositif permettant de se rapprocher le plus de la vision dominante du Bon Etat (cf. **Figure 56**), à savoir le **Scénario 3, Variante 3**, avec seuil H/G fixé à **0,76**. Tout en restant apparemment un peu plus exigeante que d'autres sur le plan du niveau de Bon Etat recherché (avec toutes les réserves déjà évoquées vis-à-vis de cette interprétation empreinte de raccourcis et d'inadéquations méthodologiques), cette proposition fait rentrer les biais assez nettement à l'intérieur de la fourchette de +/- 0,5 jugée acceptable dans le cadre de l'exercice, soit ici un biais de 0,378 pour la limite H/G et de 0,411 pour la limite G/M.

Grille finale d'évaluation TGCE proposée (Scénario 3-3 FR, 16-06-2016)

IBD 2007	REF	H/G	G/M	M/P	P/B	MIN
Region 2 - Scen. 3	17,4	16,088	13,464	9,528	5,264	1,0
Region 4 - Scen. 3	18,1	16,732	13,996	9,892	5,446	1,0
Region 6 - Scen. 3	19,1	17,652	14,756	10,412	5,706	1,0
EQR-IBD	1	0,92	0,76	0,52	0,26	0

Figure 56 : Grille finale d'évaluation des TGCE proposée par la France (Scénario 3-3, 16-06-2016)

Un dispositif calé avec ce niveau de biais a parcouru une partie du chemin qui le rapproche de la vision dominante du Bon Etat, mais ne surévalue probablement pas nos cours d'eau. D'une part, le gradient évalué par notre système ne se limite pas à l'altération par le PO4 mais à un ensemble plus composite d'altérations anthropiques ; d'autre part, les spécificités de qualité biologique de certains de nos cours d'eau, qui sont d'une échelle de taille très sensiblement inférieure à celle de certains très grands cours d'eau européens (Danube, Rhin aval...), sont bien ressorties comme une particularité propre au 2^{ème} exercice d'intercalibration.

C'est donc ce Scénario 3 – Variante 3 qui a été officiellement soumis à l'intercalibration Européenne (en date du 16-06-2016) et a servi de base nationale au rapport d'intercalibration envoyé par la Coordinatrice à l'Union Européenne en date du 21-06-2016.

En fonction du seul levier efficace possible pour se rapprocher de la vision dominante du Bon Etat, dans le cadre d'un exercice d'intercalibration d'emblée construit à partir des valeurs d'EQRs nationaux par type national de TGCE, **il n'a pas été possible** de conserver *in fine* la **même grille en EQRs** que pour les **cours d'eau plus petits**, rappelée en Figure 30, alors que c'était bien notre intention initiale, permettant de garder une logique de construction homogène de tout notre dispositif d'évaluation national ; les différences de niveau d'évaluation étant gérées par des niveaux de référence différents selon le type de cours d'eau et les zones naturelles.

Le système national d'évaluation qui en résulte se décline donc par une **nouvelle grille unique**, exprimée en EQR-IBD₂₀₀₇, de classes de qualité dédiée à l'évaluation des TGCE (cf. Figure 56), différente de celle des cours d'eau plus petits. En fonction des critères typologiques respectée dans le cadre de l'intercalibration, cette nouvelle grille devrait en principe avoir force d'application sur tous nos TGCE nationaux dont la surface intégrée de bassin versant au site d'observation dépasse le seuil européen de 10 000 km². La fin de la présente étude nationale a néanmoins envisagé différents scénarios adaptés d'application éventuellement applicables, ainsi que leurs conséquences en termes d'évaluation d'Etat et de Rapportage Européen.

C'est bien la grille d'évaluation TGCE exprimée en EQR, telle que présentée en dernière ligne de la Figure 56, qui fait référence pour le classement d'état écologique et les franchissements de seuils. Selon les types de cours d'eau et les niveaux de référence par défaut, elle se décline selon les **valeurs arrondies d'IBD** données par type national de TGCE (cf. les 3 lignes du haut de la figure).

Pour l'application du Scénario 3-3 sur le **bassin du Rhône**, la fusion de la Grille 5 (type alpin) et de la Grille 6 (type montagnard) conduit à une simplification du dispositif d'évaluation comme suit (**Figure 57**).

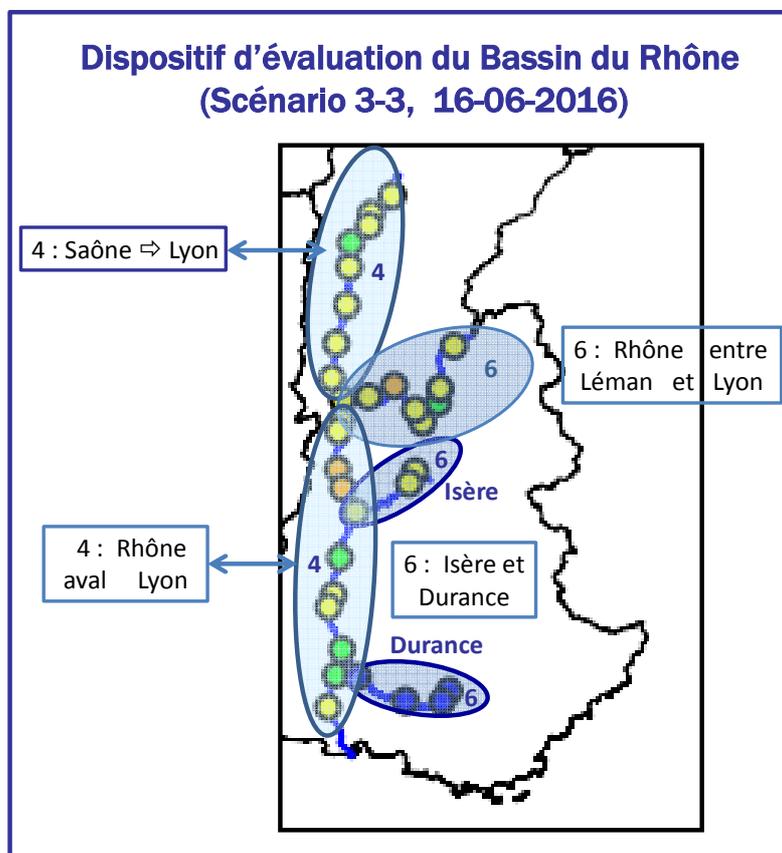


Figure 57 : Dispositif final d'évaluation TGCE à appliquer sur le bassin du Rhône (16-06-2016)

Le dispositif présenté en **Figure 57** annule et remplace donc le contenu de la Figure 35, qui présentait le dispositif à appliquer sur le Bassin du Rhône tel qu'initialement envisagé avant cette fusion des grilles 5 et 6 en une seule.

Hormis ce bassin hydrographique particulier, le contenu des autres figures 35 à 38, qui présentaient la déclinaison géographique d'application des grilles d'évaluation sur les autres grands bassins fluviaux de France, reste valide.

Au final, le rapport final de l'exercice d'intercalibration « Fit-in » auquel nous avons participé a fait l'objet d'une notification orale officielle de visa OK auprès de la Coordinatrice de l'exercice lors d'une réunion ECOSTAT organisée en Octobre 2017, la confirmation écrite étant arrivée quelques mois plus tard sous la forme d'une Décision Officielle d'Intercalibration UE datée du 12-02-2018.

6. APPUI A LA MISE EN OEUVRE DU NOUVEAU SYSTEME NATIONAL D'EVALUATION DES TGCE

Dans le cadre général de l'expertise-maillon et de l'appui à la Décision Publique apporté par Irstea au Ministère de l'Environnement et à l'ONEMA (désormais AFB), une fois la stabilisation des grilles nationales d'évaluation TGCE effectuée et l'intercalibration européenne réalisée, il restait à étudier et à consolider certaines modalités pratiques de mise en œuvre de ce nouveau dispositif, afin qu'elles puissent être décrites et faire l'objet de la publication d'un Arrêté «Evaluation» modifié officialisant le dispositif réglementaire à appliquer au niveau national.

Il restait aussi à évaluer l'impact sur le classement d'état écologique de différentes modalités pratiques de mise en œuvre envisageables pour l'application de ce nouveau dispositif. En effet, si la typologie européenne constitue un cadre général permettant d'assurer une certaine inter-comparabilité interne des données utilisées en vue d'une réalisation plus homogène d'exercices d'intercalibration européens, elle ne prend pas en charge la déclinaison complexe d'un système national d'évaluation sur différents types de TGCE ni les différents cas de figure particuliers qui sont à régler, notamment la façon d'assurer une évaluation pertinente au long d'un linéaire fluvial en fonction des problèmes de cohérence posés par la traversée successive d'HER différentes, à dispositif d'évaluation en principe différent, par des arrivées d'affluents changeant très significativement l'hydrochimie et les cortèges biologiques du cours d'eau etc...

Le présent chapitre a donc pour but de **récapituler la teneur de l'expertise-maillon et de l'appui à la Décision Publique réalisé par Irstea auprès de la Direction de l'Eau et de l'AFB entre Octobre 2017 et fin 2018-début 2019**, notamment suite à sa participation à l'exercice d'intercalibration et en vue de la révision de l'Arrêté Evaluation du 27-07-2015, en vigueur à l'époque.

Ces contributions ont notamment consisté en :

- l'envoi d'un Email Irstea d'appui à décision publique et d'une proposition de tableau d'évaluation modifié du 18-10-2017 (première proposition d'insert du dispositif TGCE dans l'Arrêté Evaluation),
- l'envoi d'un nouvel Email d'appui, accompagné d'une nouvelle proposition de tableau modifié du 30-01-2018 (idem),
- un échange téléphonique entre F. Delmas et David Carayon (Irstea, au titre de l'expertise-maillon « Diatomées des cours d'eau ») et J.P. Cabaret (Ministère - Direction de l'Eau) du 06-02-2018 (discussions, échanges, questions – réponses suscités par les propositions de modificatif d'Arrêté),
- plusieurs E-mails complémentaires d'expertise et d'envoi de supports informatifs spécifiques relatifs à l'évaluation des TGCE (notamment envoi d'un recueil de diapositives pwp...) du 06-02 au 16-02-2018),
- l'envoi d'une note d'expertise et d'appui à la Décision Publique intitulée « Evaluation de l'Etat Ecologique des Très Grands Cours d'Eau à partir du compartiment diatomique et des EQR-IBD 2007 régionalisés », version du 26-02-2018 (24 pages),
- et enfin, une nouvelle série d'échanges-modifications de fichier d'Arrêté Evaluation ayant eu lieu du dernier trimestre 2018 à la mi-Janvier 2019, afin de corriger des erreurs de frappe, de maquettage et/ou d'édition commis par le Service d'Edition du Journal Officiel dans le nouvel Arrêté Evaluation signé en date du 27 Juillet 2018 (publié au JORF du 28 Août 2018).

L'appui réalisé a concerné des aspects de rédaction-modification, dans le texte de l'Arrêté, d'inserts spécifiques à l'évaluation des TGCE, ainsi que quelques suggestions de fond et de forme relatives à la présentation et à la mise en application réglementaire du nouveau dispositif.

Quelques points restaient en suspens, notamment :

- de quelle façon évoquer les **niveaux de référence** dans le texte, sachant que ce terme recouvre plusieurs statuts de référence bien différents, qui font appel à des concepts différents (dont un niveau de référence par défaut par type de TGCE à évaluer) ; il a donc fallu statuer sur des aspects de formulation du contenu de cellules dans le tableau d'évaluation, en relation avec le statut de référence utilisé dans la case,
- traitement des **cas particuliers rencontrés par certains bassins** (exceptions typologiques, cohérence amont-aval de continuum, arrivée d'affluents à hydrochimie différente et traitement réglementaire...)
- **Critère de seuillage de la transition GCE-TGCE à retenir définitivement** pour l'application TGCE sur notre territoire national, sachant que cette limite conditionne les grilles d'évaluation à utiliser et le niveau d'évaluation produit sur un site donné. Un questionnement se pose notamment sur la façon d'articuler la typologie abiotique nationale des cours d'eau basée sur les Ordres de Strahler et la typologie UE basée sur des seuils de surface (problème de recouvrement assez important à la jonction entre ces 2 typologies).

Le contenu de la suite du présent chapitre récapitule les analyses et solutions possibles par rapport aux 3 aspects pré-évoqués. En particulier, suite à un questionnement spécifique de la Direction de l'Eau et de l'AFB concernant le 3^{ème} point (quelle typologie et quelle articulation typologique mettre en œuvre pour l'évaluation judicieuse des TGCE), il récapitule et inter-compare les résultats de 3 nouveaux scénarios possibles de mise en œuvre de l'évaluation des TGCE, par rapport aux résultats qui seraient procurés par l'application du dispositif d'évaluation contemporain (Scénario 1 : application dans l'état de l'Arrêté du 27 Juillet 2015).

Les résultats statistiques procurés par l'application de ces 3 scénarios de mise en œuvre de l'évaluation y sont comparés et commentés, et quelques suggestions et/ou recommandations opérationnelles, du point de vue des experts-maillon, sont apportées pour aider la prise de décision des tutelles vis-à-vis des choix finaux relatifs à la mise en application du nouveau dispositif.

En fin de chapitre il est fait le point sur l'Appui à Décision Publique et sur l'expertise-maillon mise au service de l'évolution du cadre réglementaire national pour l'évaluation des TGCE, ainsi que du document de transfert destiné à venir en appuyer la mise en œuvre (actualisation 2019 du Guide REEE).

6.1. Critères de définition des TGCE et étude des conséquences sur l'évaluation des TGCE de France

6.1.1. Aspects généraux et contextuels

L'investigation du gisement de données national sous-jacent à la mise au point d'un nouveau dispositif d'évaluation « TGCE » s'est confrontée en premier lieu à un problème de **définition** de ce que sont ou doivent être les **Très Grands Cours d'Eau**. En effet, le bornage des requêtes nécessaires pour extraire les données sous la BDD PANDORE et pour entreprendre le chantier national de mise au point d'un nouveau dispositif d'évaluation des TGCE a déjà été soumis à ce questionnement, sachant qu'il valait mieux extraire directement toutes les données éventuellement nécessaires pour réaliser ce travail (qui peut le plus peut le moins...) que procéder à des extractions éventuellement trop restreintes sur lesquelles il faudrait revenir plus tard sous forte contrainte de temps.

Un autre volet de travail à réaliser à peu près en même temps a aussi pesé en faveur de cette logique de raisonnement : comme le GIg sollicitait la démonstration de nos relations nationales pression-impact à l'appui de l'intercalibration de la méthode proposée (qui sont toujours affaiblies sur de très grands cours d'eau par rapport à des hydrosystèmes plus petits...), il valait mieux tenter de bâtir les relations les mieux assises possibles, ce qui passe par un jeu de données suffisamment étoffé à la fois sur le plan du nombre de relevés pris en compte et de la diversité des situations couvertes. En effet, une meilleure assise en termes de nombre de stations et de dates, intégrant les variabilités attachées, contribue à étendre et à raffiner la représentation des gradients anthropiques et de leurs conséquences en termes de réponses diatomiques, notamment par l'extension du domaine des possibles, en fonction des lieux et des dates, sur le gradient d'enrichissement en nutriments contenu dans le jeu de données d'étude (depuis des sites naturels jusqu'aux situations les plus fortement polluées).

Sur le plan du repérage typologique des TGCE et dans l'absolu, différents critères, dont l'un n'est pas plus légitime que l'autre, peuvent contribuer à définir ces objets hydrographiques selon leur taille. *A priori*, il serait possible de s'appuyer sur : la distance entre la source et le point d'observation ; la surface cumulée de bassin versant drainé au point d'observation ; l'ordre de Strahler, différents marqueurs hydromorphologiques de la largeur du cours d'eau (largeur du lit mineur, largeur plein bord, etc...) ; sa section mouillée dans différents cas de figure liés au régime (idem) ; le débit nominal au point d'observation, etc...

Tous ces critères, pris séparément ou considérés de façon plus intégrée, peuvent légitimement être utilisés pour représenter un gradient de taille des hydrosystèmes « cours d'eau » sur des bases consistantes et pour en établir une typologie basée sur leur importance.

La typologie nationale sous-jacente à l'application des différents Arrêtés de Surveillance et d'Evaluation parus depuis 2005, année de mise en place du premier dispositif national de surveillance dans le cadre de la DCE, classe les types de cours d'eau selon leur contexte géoclimatique d'appartenance (HERs) et selon des critères de taille basée sur les Ordres de Strahler (depuis les plus petits cours d'eau permanents jusqu'aux TGCE). Sur le plan de la taille des cours d'eau, cette classification proposée au début des années 2000 par l'équipe du Laboratoire d'Hydrologie Quantitative (Irstea Lyon / LHQ) s'est calée sur les **Ordres de Strahler** contenus dans la BD Carthage, vision indirecte de la taille de l'hydrosystème donnée par le niveau cumulé d'ordres de confluences depuis la source du drain principal jusqu'au site d'observation.

Il a été remarqué que cette typologie, si elle classe relativement bien les systèmes selon des conditions géologiques, édaphiques et climatiques par zone d'appartenance, pouvait donner une classification assez bruitée, à l'échelle nationale, par rapport à d'autres critères tout aussi légitimes que l'ordre pour rendre compte d'une certaine hiérarchie de taille des hydrosystèmes (ex : longueur du cours d'eau depuis sa source, surface intégrée de bassin versant au site d'observation, etc...). En effet, il se trouve d'une part que le coefficient de drainage d'un bassin versant au km² est assez sensiblement variable en fonction du milieu naturel, des pentes, de la nature et de la perméabilité des substrats géologiques, du régime de pluies dans la région naturelle considérée etc... Or le coefficient de drainage du bassin versant et les caractéristiques morphologiques du réseau des linéaires de surface vont influencer directement sur la relation taille-ordre.

D'autre part, les Ordres de Strahler ont fait l'objet d'un calage initial selon des principes qui ont pu différer selon les grands bassins hydrographiques. En effet, chaque Agence de Bassin a eu à construire, à un moment donné, un SIG de représentation et de hiérarchisation des hydrosystèmes de son propre bassin hydrographique (BD Carthage, dérivée de la BD Carto de l'IGN). Les contextes naturels et les façons de faire au sein de chaque Agence présentant des différences, chacune d'entre elle a pu définir les ordres des cours d'eau sur son territoire avec l'incorporation d'une part plus ou moins importante d'effet-opérateur ; un aspect particulièrement sensible à ce niveau pouvant être une prise en compte différente et éventuellement décalée du début des très petits cours d'eau permanents.

Ainsi, entre autres hétérogénéités inter-bassins plus ou moins marquantes, le Bassin Loire-Bretagne est celui où un ordre donné de cours d'eau est en général atteint pour des cours d'eau sensiblement plus petits que sur tous les autres grands bassins hydrographiques de France. D'autres hétérogénéités de typologie physique, bien que moins générales, ont aussi pu être relevées dans d'autres bassins, au moins dans certains contextes naturels bien spécifiques (cas notamment de l'Aude dont le bassin versant assez restreint, adossé au Pied-Mont Pyrénéen et au Massif des Cévennes subit un régime pluviométrique sous influence Cévenole avant de se jeter en Méditerranée ; cas de la Maine, réseau hydrographique de plaine à hydrodynamique nettement plus atténuée...), conduisant à quelques situations fortement atypiques et à des exceptions notables dans les 2 sens dans la relation entre Ordre de Strahler et surface intégrée de bassin versant à la station de réseau.

Dans le cadre de cette étude faisant partie de la Fiche-Action AFB-Irstea N°23 (2016-2018), nous avons à mener 2 chantiers de front :

- un travail national autocentré d'investigation de données TGCE en vue de proposer un dispositif national d'évaluation de cette catégorie de cours d'eau,
- la représentation nationale et la fourniture de données nationales dans le cadre du nouvel exercice d'intercalibration européenne TGCE qui s'est rapidement mis en place au 2^{ème} semestre 2015.

Cette double implication nous a conduits à devoir composer en permanence avec une **différence de pratique** entre notre **niveau national** et la définition typologique qu'il nous était demandé de respecter **dans le cadre de l'exercice d'intercalibration européenne**.

Dans le cadre UE, concernant la minéralisation de l'eau, il nous a été demandé de vérifier le respect d'une typologie hydrochimique européenne basée sur des seuils d'alcalinité exprimé en milli-équivalents de carbonates, descripteur non-utilisé en France et dont la reconstitution et la vérification de conformité ont nécessité un important post-traitement des données physico-chimiques.

De même, le critère physique de taille utilisé au niveau européen pour cadrer ces exercices et leur assurer une relative homogénéité interne (exigence renforcée du fait que certains jeux de données nationaux contiennent des données relatives à des cours d'eau pluri-nationaux comme le Rhin et le Danube, évalués successivement par plusieurs pays à l'échelle de l'Europe), s'appuie sur une surface minimale cumulée de bassin versant à la station d'observation **supérieure ou égale à 10 000 km²**. Le critère national de typification des TGCE se basant pour sa part sur les Ordres de Strahler, il intervient donc des hétérogénéités de classification, dans certains cas assez importantes, qui posent question pour l'application réglementaire de seuils.

Concernant la fourniture du jeu de données d'intercalibration, il a fallu respecter le seuil de taille imposé dans le cadre UE, ce qui a imposé la réalisation d'un travail spécifique visant à situer les surfaces intégrées de bassins versants aux sites de nos stations de réseaux.

Au moment du démarrage de ce chantier (entre Juin et Septembre 2015), ces données de surface intégrée au site de réseau étaient encore rarement disponibles dans la bancarisation nationale Sandre-SIE, et disponible de façon approchée et parfois incomplète via les référentiels internes Irstea Pandore : 1) données Dahu 2004-2005 (surface donnée à la station) ; 2) données Syrah 2008-2009 (surface intégrée donnée à l'aval du tronçon de masse d'eau concerné).

Le seuil de typologie édicté pour la réalisation d'exercices européens d'intercalibration a pour but principal de garantir une certaine homogénéité et comparabilité *a priori* des jeux de données nationaux fournis dans le cadre de ces exercices collectifs. Il se base sur un indicateur de taille objectif et assez facile à renseigner par les Pays-Membres. Son objectif premier n'est pas forcément de venir imposer des seuils européens dans le cadre de l'élaboration et de l'application réglementaire des systèmes nationaux d'évaluation, beaucoup plus proches des réalités biogéographiques de terrain de chaque Pays-Membre.

Si cette différence de pratique entre notre niveau national et l'UE montre bien que 2 critères au moins, entre autres tout aussi légitimes, peuvent être utilisés pour typifier les TGCE, cela pose aussi question, en retour, sur la façon non contestable et opposable d'appliquer ensuite au niveau national les résultats d'un exercice d'intercalibration réalisé selon certaines modalités précises et visé OK par l'Union Européenne.

Ces éléments de contexte interrogent donc sur le critère le plus pertinent qu'il conviendrait de prendre en compte pour venir en appui au seuillage des TGCE dans notre système national d'évaluation. Sans que des interlocuteurs scientifiques soient eux-mêmes en position de trancher cette question, qui touche à l'application d'une politique d'évaluation dans différents cadres réglementaires (l'un national, l'autre européen, dans le processus d'application nationale de la DCE), il était clair dès le départ que baser ce type de classification permettant d'approcher une taille de cours d'eau et de son réseau hydrographique selon ces 2 approches conceptuellement bien différentes n'allait pas produire une vision parfaitement concordante en sortie.

Est-ce une conséquence de pratiques différentes de chaque bassin pour élaborer ses propres couches d'information BD Carthage, ou de raisons hydromorphologiques très particulières à certains bassins, ou d'un composite des 2 ? Le fait est que la limite inférieure de définition des TGCE dans la typologie nationale (type TG) correspond avec des surfaces intégrées de bassin versant très variables et pouvant s'éloigner assez fortement du TGCE-type tel qu'utilisé dans la typologie d'intercalibration. Entre autres incohérences sur la zone de jonction entre les 2 typologies, le plus petit fleuve français qualifié de TGCE selon les ordres de Strahler est l'Aude à partir de Carcassonne (2 768 km²), ce qui constitue un vrai outlier hydromorphologique et typologique qu'il convient de rectifier. En effet, si c'est un cours d'eau assez profond (non ou difficilement wadeable) dont les échantillonnages d'invertébrés peuvent poser problème selon la méthodologie classique, c'est un cours d'eau exceptionnellement petit, beaucoup trop éloigné des critères de taille habituellement attachés à la catégorie des TGCE. La typologie nationale est aussi porteuse d'un flou sur des tailles nettement plus grandes, s'approchant ou dépassant assez largement ce qui serait considéré comme TGCE dans d'autres types de référentiels basés sur les surfaces. Beaucoup de cours d'eau encore classifiés en grands cours d'eau (GCE) et même, plus rarement, en moyens cours d'eau (MCE) selon notre typologie nationale dépassent déjà largement le seuil de surface de 10 000 km², dont plusieurs gros affluents de la Seine et la Seine elle-même (entre 12 000 et 13 000 km²) ; l'exemple le plus extrême étant la Maine à Bouchemaine, encore qualifiée de GCE alors que sa surface cumulée de BV dépasse les 23 000 km² à cette station.

La typologie nationale comporte donc un chevauchement de classes de taille qui peut couvrir un rapport de plus de 1 à 8 entre les extrêmes typologiques si l'on tient compte de l'Aude, de près de 1 à 4 si l'on ne tient pas compte de cet extrême au regard de cette typologie hydromorphologique.

Donc, même si les surfaces de bassins versants intégrés actuellement disponibles sur les sites de réseaux de surveillance des cours d'eau français sont données en intégrant un certain niveau d'approximation (selon la résolution du MNT utilisé, selon sa précision altimétrique, selon la programmation de l'algorithme utilisé pour modéliser le tracé de drainage depuis les lignes de crête du BV jusqu'à la station de réseau...), le critère de surface intégrée de bassin versant à la station de réseau semble un indicateur plus précis et plus reproductible, d'un bassin hydrographique à l'autre, qu'un critère typologique de taille basé sur les Ordres de Strahler.

Il n'est donc pas inutile d'examiner et de diagnostiquer l'incidence sur l'Evaluation d'Etat Ecologique des TGCE d'un seuil établi par rapport à des **surfaces intégrées**, dont le résultat est à coup sûr plus précis et certainement plus inter-comparable d'un bassin à l'autre, par rapport à ce seuil TGCE basé sur les ordres de Strahler et à l'imprécision sur la classification de taille qui en résulte (nombreux chevauchements inter-types de tailles entre MCE, GCE et TGCE).

Dans les scénarios d'étude des résultats nationaux d'évaluation des TGCE, nous avons donc comparé une évaluation des TGCE basée sur la **typologie nationale** et une évaluation des TGCE définis selon la typologie d'intercalibration européenne (cours d'eau au bassin versant intégré à **partir de 10 000 km²**).

En fonction du contexte de notre territoire national, il est aussi apparu utile d'étudier l'éventualité d'un seuil de surface un peu abaissé (**8000 km²**), susceptible de constituer un bon compromis national pour plusieurs raisons :

- Pour l'instant, l'hétérogénéité du critère « ordre de Strahler » selon les grands bassins hydrographiques conduisait déjà à une classification de certains cours d'eau français comme TGCE à partir de surfaces cumulées de bassin versant de 6 300 km² si l'on excepte l'Aude (soit environ 40 % en-dessous du seuil typologique européen utilisé dans le cadre de l'intercalibration) ;

- La plupart des grands affluents des 5 fleuves Français avant leur confluence avec le drain principal sont des hydrosystèmes de taille déjà conséquente, pour lesquels le système actuel d'évaluation diatomique en vigueur en France a été calé de façon assez approximative du fait de la lacune nationale en situations de référence représentatives d'aussi grands cours d'eau de plaine. A ce niveau de taille, on se situe déjà largement dans la même problématique que celle qui a conduit à retarder la mise au point d'un système national d'évaluation TGCE (pression anthropique plus ou moins forte mais omniprésente, d'origine agricole ou liée à une succession d'implantations urbaines au long du drain considéré, et absence de références vraies valides pour cette échelle de taille). Sur notre territoire, pas mal d'affluents principaux de nos 5 grands fleuves font plus de 8 000 km² et ont au moins une station de surveillance implantée avant leur confluence avec le drain principal. Ce sont déjà des cours d'eau de plaine de taille importante avec des cortèges biologiques indiscutablement marqués par cette situation relativement aval dans le continuum fluvial (e.g. river continuum concept, *Vannote et al*, 1980), qu'il serait judicieux d'évaluer dans un système homogène en bonne continuité avec le drain fluvial situé juste à l'aval. Or, avec les assez grandes hétérogénéités de surface intégrée attachées à la classification nationale basée sur les ordres (larges recoupements entre classifications (MCE), GCE, TGCE de 6 000 km² à plus de 23 000 km²), cette homogénéité n'est pas garantie et peut mener à un traitement d'évaluation différent pour des masses d'eau d'égale importance selon les bassins hydrographiques.

- La typologie de taille européenne a été raisonnée à sa large échelle de territoire, dans l'objectif de garantir une relative homogénéité des données proposées par les pays-membres pour les exercices d'intercalibration. L'homogénéité des données de ce point de vue, qui était le but de cette typologie, est relativement assurée pour les cours d'eau pas trop éloignés du seuil inférieur. Cependant, on peut aussi objecter qu'elle devient nettement plus discutable lorsque les jeux de données nationaux intègrent aussi de très conséquents fleuves inter-frontaliers dont certains (Rhin : jusqu'à 198 000 km², soit près de 20 fois le seuil, et Danube : plus de 800 000 km², soit plus de 80 fois le seuil) sont d'un ordre de taille exceptionnellement grand par rapport à la plupart des autres TGCE de notre pays et de nos pays voisins.

- Dans leurs systèmes nationaux d'évaluation actuellement en vigueur, nos voisins ont décliné leur propre typologie nationale, dont nous ne connaissons pas la déclinaison précise en détail (elle est rarement élaborée comme une déclinaison de la typologie d'intercalibration, puisqu'elle pré-existait au lancement de ces exercices). Certains de nos Pays-Membres voisins peuvent donc évaluer aussi en TGCE certains cours d'eau de taille comparable à nos grands affluents de plaine.

Une certaine marge de souplesse reste donc envisageable pour l'adaptation typologique d'évaluation en bonne adéquation avec les caractéristiques de chaque réseau hydrographique national, ce qui est du ressort de chaque Etat-Membre. Ce qui doit être garanti est l'évaluation du Bon Etat à un niveau comparable (cf. rentrer dans les biais jugés acceptables dans les exercices d'intercalibration...), vérifiée dans le cadre de tels exercices collectifs portant sur un jeu de données comparable. La mise en correspondance entre types nationaux et typologie d'intercalibration, souvent de plus large échelle et moins raffinée que les typologies nationales, entraîne souvent la nécessité de sélection d'un sous-jeu de données incomplet par rapport à des types nationaux, mais correspondant aux critères typologiques exacts demandés dans le cadre de l'exercice.

6.1.2. Descriptif de 4 scénarios envisageables pour la mise en œuvre de l'évaluation des TGCE

Dans le but de venir éclairer la Décision Publique sur les modalités les plus souhaitables de mise en application nationale du nouveau dispositif d'évaluation des TGCE, récemment intercalibré UE, nous avons étudié et mis en comparaison les résultats d'évaluation des TGCE qui seraient obtenus par 4 scénarios d'évaluation basés sur 3 critères différents de définition du seuil typologique de taille « TGCE » : 1) classement TGCE selon la typologie nationale, 2) classement TGCE à partir d'un bassin versant intégré de 10 000 km², 3) classement TGCE à partir d'un bassin versant intégré de 8000 km².

Ces scénarios ont été calculés sur la base de **671 opérations de contrôle biologique** (relevés diatomiques) réalisées sur un créneau temporel de 8 ans (2007-2013) sur **127 sites** de surveillance de cours d'eau répartis au long de **22 cours d'eau différents** (sites de réseaux et cours d'eau susceptibles d'être évalués en tant que TGCE selon le critère de sélection de taille utilisé).

- **Scénario 1** : Il s'agit du **scénario d'évaluation contemporain** (= situation actuelle à Décembre 2017-Janvier 2018, avant passage dans l'application du futur Arrêté Evaluation modifié), qui résulterait de l'application dans l'état du contenu de l'Arrêté du 27 Juillet 2015 (unique adaptation, commune aux 4 scénarios : Aude aval repassée en GCE pour cause de trop fort décalage typologique). Il convient de faire ressortir que ce scénario ne peut pas servir de référence autre qu'indicative, puisqu'il n'évalue pas les Très Grands Cours d'Eau Alpains (Rhin et Rhône avant son passage aval en domaine calcaire méditerranéen) et qu'il évalue les autres TGCE selon une grille historique non-intercalibrée et trop sévère.
- **Scénario 2** : Ce scénario **complètement basé sur la typologie nationale** des cours d'eau (basée sur les ordres de Strahler) combine 2 dispositifs d'évaluation : 1) le dispositif de l'Arrêté du 27-07-2015 pour **tous les cours d'eau plus petits que TGCE et TGCA selon la typologie nationale**, et 2) le nouveau dispositif proposé et intercalibré pour les TGCE, **appliqué sur les Très Grands Cours d'Eau déterminés selon notre typologie nationale**. A noter que ce dispositif, qui génère les hétérogénéités d'évaluation de taille pré-évoquées, peut conduire à des différences de traitement d'évaluation pour des cours d'eau de même surface intégrée, en fonction du grand bassin hydrographique Français dans lequel on se situe.
- **Scénario 3** : Ce scénario est construit sur l'application du **seuil typologique de surface de 10 000 km²** qui a permis de sélectionner les sites TGCE dans le cadre de l'intercalibration européenne « Large Rivers ». Ce scénario d'évaluation est composite, à savoir qu'il évalue tous les cours d'eau plus petits que 10 000 km² selon les grilles de l'Arrêté du 27-07-2015, et tous les cours d'eau atteignant ou dépassant ce seuil TGCE de 10 000 km² selon les nouvelles grilles d'évaluation du dispositif Français TGCE intercalibré.
- **Scénario 4** : Ce scénario est construit sur l'application d'un **seuil typologique de surface de 8 000 km²**, seuil un peu plus bas que celui ayant servi pour l'intercalibration Européenne sur un jeu de données de large échelle, mais possiblement en meilleure adéquation avec les caractéristiques de notre réseau hydrographique national et notamment, du niveau des confluences entre nos 5 grands fleuves français et leurs affluents principaux. En effet, le plus souvent, ceux-ci s'écoulent déjà depuis longtemps dans le domaine de plaine et ne peuvent légitimement plus être directement comparés, même à ce seuil de taille, aux situations de référence réelles calculées pour la Région Diatomique donnée, assises sur des cours d'eau nettement plus petits.

6.1.3. Assise et résultats bruts des 4 scénarios d'évaluation

La mise en œuvre des 4 scénarios et l'examen de leurs résultats a fait l'objet d'un travail d'intercomparaison réalisé en Janvier-Février 2018. Un fichier Excel contenant le jeu de données d'étude et les résultats produits par les 4 scénarios, depuis le relevé individuel jusqu'aux résultats statistiques d'évaluation à l'échelle nationale pour les TGCE, a été transmis à la DEB et à l'AFB en même temps que la note d'appui à Décision Publique du 26-02-2018, dont le but était de venir appuyer des prises de décision pouvant avoir des répercussions directes sur certains contenus du nouvel Arrêté Evaluation (tableur Excel « Table comparaison_4 scenarios_TGCE.xls »).

Pour information, au cas où il serait utile de ré-examiner le contenu de ce tableur, la feuille située le plus en fond de classeur intitulée « table globale » contient les informations essentielles sur les sites et relevés, y compris les notes d'IBD acquise au site à la date du relevé et d'EQR calculées selon les scénarios 1 (situation actuelle) et 3 (seuil TGCE calé à 8 000 km²), correspondant à la sélection exhaustive des sites et relevés candidats à une évaluation selon le statut TGCE.

S'ensuivent 4 feuilles correspondant à l'évaluation et à la classe de qualité obtenue selon chacun des 4 scénarios tels que formalisé ci-dessus, puis une feuille Synthèse_Scénarios qui liste en parallèle les 4 résultats d'évaluation pour chaque Opécont élémentaire.

La feuille suivante donne l'exploitation statistique de ces classifications d'état au relevé selon les 4 scénarios, et agrège aussi les résultats selon un tableau récapitulatif permettant la comparaison directe de leurs résultats.

Enfin, la feuille « Description » détaille la signification des sigles de variables présentées dans les tableaux et résume le principe d'élaboration de chacun des 4 scénarios d'évaluation appliqués.

Les résultats de chaque scénario d'évaluation étudié sont présentés dans les pages suivantes.

Les 2 premiers scénarios d'évaluation produits ci-dessous récapitulent les résultats d'évaluation obtenus au relevé selon un système utilisant la **typologie nationale** comme base de définition des TGCE.

Le Scénario 1 présenté dans le **Tableau 5** ci-dessous rend compte de la situation d'évaluation en vigueur au moment de l'étude, qui résulte de l'application dans l'état du contenu de l'Arrêté du 27 Juillet 2015 sur le jeu de relevés diatomiques sélectionné. Le système d'évaluation qu'il contient ne découle pas d'une réflexion spécifique « Très Grands Cours d'Eau », mais d'une simple propagation vers l'aval du contenu de grilles d'évaluation dédiées à de plus petits cours d'eau.

Les Très Grands Cours d'Eau Alpains (TTGA) n'y sont pas évaluable (pas de contenu réglementaire en vigueur pour les cases qui les concernent), d'où un gap de 152 relevés non évalués par rapport aux 671 relevés du jeu de données disponible.

Tableau 5 : Résultats d'évaluation diatomique au relevé selon classification TGCE basée sur la typologie nationale : Scénario 1 : Situation actuelle (cf. Arrêté du 27-07-2015 actuellement en vigueur)

Scénario 1					
Classe de qualité	Effectif total évalué	Pourcentage	Pourcentage (sans les NA)	Effectif de TGCE évalués (sans les NA)	Pourcentage TGCE évalués
Très bon	44	6,56	8,48	33	7,69
Bon	113	16,84	21,77	77	17,95
Moyen	287	42,77	55,30	259	60,37
Médiocre	74	11,03	14,26	59	13,75
Mauvais	1	0,15	0,19	1	0,23
Total	519	77,35	100,00	429	0,00

NA	152	22,65	Selon Scénario 1 : 90 relevés sur 671 = plus petits que TG		
----	-----	-------	--	--	--

Ce premier système, présenté seulement à titre indicatif, a simplement permis de produire une statistique d'état basée sur la pratique réglementaire actuelle, très perfectible. Il n'a pas de légitimité réelle pour évaluer de façon judicieuse les TGCE à l'avenir. En effet, les grilles contenues dans l'Arrêté du 27-07-2015 ont été intercalibrées pour des hydrosystèmes plus petits (jusqu'aux types nationaux GCE), mais n'ont pas été conçues ni intercalibrées pour évaluer de façon judicieuse les très grands cours d'eau. C'est d'ailleurs la raison pour laquelle il a été travaillé, entre 2015 et 2017, à l'élaboration et à l'intercalibration au niveau européen d'un nouveau dispositif national spécifiquement dédié à l'évaluation de tels cours d'eau.

Par rapport à ce **Scénario 1**, les travaux comparatifs inter-Etats-Membres menés à l'occasion de ce chantier européen sur les TGCE ont bien révélé qu'il convenait d'adopter, sur les grands hydrosystèmes nationaux, des niveaux de référence par défaut assez sensiblement plus bas que les niveaux de référence vraie calculés sur de plus petits cours d'eau. Par contrecoup immédiat, cela conduit logiquement à devoir repousser mécaniquement vers des valeurs un peu plus basses, en notes d'IBD 2007, les limites inférieures de basculement des classes de qualité et en particulier la limite inférieure de la Classe de Bon Etat.

Comme l'exercice collectif d'intercalibration européenne a été réalisé directement sur l'intercomparaison des résultats d'évaluation exprimés en EQR, sans examen spécifique des niveaux de référence par défaut choisis par chaque Etat-membre par type national de TGCE, nous avons de plus été contraints en cours d'exercice d'abaisser un peu les seuils de la grille d'EQR proposée pour l'évaluation des TGCE par rapport à la grille d'EQR utilisée pour tous les types de plus petits cours d'eau (0,92 au lieu de 0,94 pour la limite TBE-BE et 0,76 au lieu de 0,78 pour la limite BE-EM), afin de pouvoir se rapprocher efficacement de la vision dominante du Bon Etat qui est ressortie de ces exercices collectifs.

Ces éléments de comparaison avec les autres Pays-Membres participants et les ajustements à la baisse qui ont été nécessaires enlèvent donc tout intérêt autre qu'indicatif au scénario contemporain (Scénario 1), qui ne peut pas servir de référence valide pour l'évaluation des TGCE.

Le **Scénario 2** présenté au Tableau 6 page suivante présente la situation d'évaluation composite qui entrera en vigueur avec l'incorporation du nouveau dispositif d'évaluation TGCE 2018, dans le cas où on continue d'utiliser la **typologie nationale** comme critère d'identification des TGCE. Pour les cours d'eau non classifiés TGCE dans cette typologie (GCE et, de façon plus marginale, MCE), l'évaluation utilise les grilles dédiées aux plus petits cours d'eau (cf. contenu de l'Arrêté du 27 Juillet 2015), avec un petit nombre d'amendements nécessaires apportés, dans l'actuelle révision d'Arrêté, pour l'évaluation de l'Aude (trop petite et atypique, retirée des TGCE), la Meuse, l'Isère aval.

Tableau 6 : Résultats d'évaluation diatomique au relevé des TGCE définis selon la typologie nationale : Scénario 2 : Application des nouvelles grilles 2018 sur TGCE, règles 2015 sur cours d'eau non-TGCE

Scénario 2				
Classe de qualité	Effectif total évalué	Pourcentage	Effectif de TGCE évalués	Pourcentage TGCE évalués
Très bon	56	8,35	45	7,75
Bon	209	31,15	173	29,78
Moyen	339	50,52	311	53,53
Médiocre	65	9,69	50	8,61
Mauvais	2	0,30	2	0,34
Total	671	100,00	581	100,00

Selon Scénario 2 : 90 relevés sur 671 = plus petits que TG

Il convient de mentionner que, sur les 671 relevés sélectionnés, 90 non typifiés TGCE selon la typologie nationale continuent d'être évalués avec des grilles dédiées aux plus petits cours d'eau, en fonction du biotype de rattachement couplé à cette typologie nationale. L'Aude ayant été enlevée d'emblée du jeu de données TGCE (exception typologique la plus importante cumulant seulement 2 600 km² de surface intégrée de bassin versant à Carcassonne), un petit nombre de TGCE définis selon cette typologie ont des surfaces intégrées de BV comprises entre 6 000 et 8 000 km². *A contrario*, plusieurs d'entre eux, en particulier dans le Bassin Parisien (dont la Seine elle-même), sont encore classifiés dans le type GCE jusqu'à des surfaces dépassant les 12 000 à 13 000 km², et même jusqu'à 23 000 km² dans le cas extrême de la Maine (affluent de la Loire vers Angers qui cumule les bassins versants de la Sarthe et de la Mayenne). Ce sont ces situations qui conduisent à ce que 90 relevés restent évalués selon l'ancien système.

Les résultats d'évaluation des scénarios seront inter-comparés et commentés plus en détail en dernière partie de cette note. Cependant, il peut déjà être relevé à ce stade que, selon cette typologie basée sur les Ordres de Strahler, dans un contexte national où il est manifestement intervenu un effet-opérateur sensible entre bassins hydrographiques, une telle hétérogénéité de taille de systèmes selon cette classification conduit à un chevauchement important, dans les 2 sens, entre plusieurs catégories de tailles (en particulier entre GCE et TGCE, mais aussi avec quelques MCE).

Cet état de fait clairement constaté, auquel il est même fait référence explicite dans l'Arrêté Evaluation de Juillet 2015 (cf. décalage typologique du bassin LB avec les autres), ouvre la porte à un traitement d'évaluation assez sensiblement différent, pour des cours d'eau du même ordre de surface, d'un bassin hydrographique à l'autre.

En effet, au-delà du rattachement à une grille d'évaluation biotypologique ou à une autre en fonction de la flore naturelle présente dans une région naturelle et/ou une HER donnée (application d'une grille biotypologique N° x ou y), le rattachement à la catégorie TGCE ou à une catégorie de plus petits cours d'eau va mécaniquement faire basculer l'évaluation vers un système de grilles d'EQR calé différemment, à la fois sur le plan du niveau de référence pour le calcul des EQR et des valeurs d'EQR de basculement inter-classes de qualité (les grilles pour les plus petits cours d'eau étant plus sévères à ces 2 niveaux que les grilles TGCE).

Il paraît donc souhaitable de veiller à limiter les hétérogénéités de classification de taille des hydrosystèmes, en particulier dans ce secteur de chevauchement typologique, afin de ne pas faire intervenir un niveau trop élevé d'aléa dans une affectation de grille d'évaluation (dans l'absolu, mais surtout du point de vue de l'équité du niveau d'évaluation en inter-bassins hydrographiques).

Les 2 autres scénarios étudiés s'appuient donc sur un critère utilisé par l'Union Européenne pour typifier les cours d'eau du point de vue de leur taille, à savoir se baser sur des **surfaces intégrées de bassin versant à la station d'évaluation** pour déterminer la limite entre plus petits cours d'eau et TGCE. Ce critère est *a priori* tout aussi pertinent et légitime qu'un autre pour établir une typologie de cours d'eau basée sur la taille. Il présente en outre l'avantage d'être relativement objectif et de limiter l'imprécision des évaluations, qui devient de plus en plus réduite en fonction de la précision actuelle des outils d'interprétation géographique type SIG.

De ce fait, les surfaces intégrées de bassins versants sont couramment utilisées pour définir des types européens de cours d'eau dans le cadre de l'intercalibration européenne, afin de garantir une plus grande homogénéité interne du jeu de données constitué à partir des données de différents Pays-Membres.

En effet, outre leur utilisation dans le cadre de l'intercalibration TGCE, elles sont intervenues aussi dans la définition de la plupart des types européens de cours d'eau à intercalibrer, la typologie faisant intervenir à la fois une classification de surface et des éléments sur la minéralisation des eaux (types CB 1 à 6, types MED 1 à 4 etc...).

Bien sûr, l'évaluation ou la modélisation de surface donnée au site incorpore forcément un certain degré d'imprécision, éventuellement variable selon le type d'environnement physique et les outils utilisés (notamment MNT, sa résolution, sa précision altitudinale, l'algorithme de modélisation utilisé pour rejoindre les lignes de crête d'interfluve au site particulier de la station de réseau...). Mais, sauf exception liée à une erreur de détournement ou à une exception morphologique, elle est en général contenue en-dessous de 1 à 2 %. Elle est donc loin d'entraîner un chevauchement typologique sur des écarts de taille pouvant varier de 1 à 5 (500 %), voire 1 à 8 (800%) pour des extrêmes typologiques, comme cela peut être observé avec la typologie nationale basée sur les Ordres de Strahler.

Le Scénario 3 présenté au **Tableau 7** ci-dessous s'appuie sur l'utilisation complète de la sélection de sites candidats-TGCE extraite de Pandore au seuil de requête de **8 000 km²**, et évalue selon les nouvelles grilles TGCE proposées les **671 relevés diatomiques** obtenus sur 22 fleuves et affluents dépassant cette taille.

Tableau 7 : Résultats d'évaluation diatomique au relevé selon le **Scénario 3** (application des nouvelles grilles d'évaluation sur les TGCE définis au seuil de surface > 8 000 km²).

Scénario 3				
Classe de qualité	Effectif total évalué	Pourcentage	Effectif de TGCE évalués	Pourcentage TGCE évalués
Très bon	57	8,49	57	8,49
Bon	214	31,89	214	31,89
Moyen	336	50,07	336	50,07
Médiocre	62	9,24	62	9,24
Mauvais	2	0,30	2	0,30
Total	671	100,00	671	100,00

Selon Scénario 3 : les 671 = tous TGCE (seuil 8 000 km²)

Le seuil de surface utilisé ici est un peu plus bas que le seuil européen de 10 000 km² utilisé pour cadrer la réalisation des exercices d'intercalibration TGCE. Cependant, il ne recoupe que très modérément le chevauchement de classes typologiques du système actuellement utilisé à notre niveau national (typologie basée sur les ordres). Sans s'attarder sur des cours d'eau à classement trop exceptionnel comme l'Aude, la typologie en vigueur classait déjà en TGCE plusieurs cours d'eau à partir de surfaces intégrées comprises entre 6000 et 7000 km². Il ne classait pas encore en TGCE des cours d'eau dont plusieurs jouxtent les 13 000 km² et même 23 000 km² dans un cas (la Maine, GCE).

Le seuil de 8 000 km², plus précis et générant moins de variabilité et de chevauchement inter-types que la typologie nationale actuellement utilisée, aurait l'avantage d'évaluer avec le même type de grilles de qualité, donc de façon plus homogène et comparable, dans un réseau hydrographique donné mais aussi et surtout, entre grands bassins hydrographiques français, les affluents fluviaux assez importants avant leur confluence avec le fleuve principal. Ce seuil de taille pourrait donc bien correspondre aux caractéristiques des hydrosystèmes fluviaux de notre territoire, en vue d'une application nationale d'évaluation adaptée à leurs spécificités.

A titre de relativisation de la portée du seuil de taille européen de 10 000 km², les exercices d'intercalibration sont cadrés avec des critères typologiques assez grossiers et propres à la large échelle Européenne. Par exemple, des fleuves très importants comme le Rhin aval (X 20 fois ce seuil) et même le Danube aval (X 80 fois ce seuil), d'un ordre de taille **beaucoup** plus important que ce qu'on peut qualifier de grands hydrosystèmes au niveau du territoire Français, ont participé à ces exercices. Si, au sein d'un même exercice d'intercalibration, des seuils propres à l'exercice doivent être appliqués assez rigoureusement afin de garantir une certaine homogénéité relative du jeu de données européen agrégé à partir de plusieurs jeux de données nationaux, condition utile pour évaluer sur une base la plus comparable possible les performances d'évaluation des différentes méthodes nationales, leur vocation première n'est pas de se substituer aux systèmes nationaux de classification et d'évaluation, et ils ont seulement une valeur indicative à ce niveau.

Dans le cadre de la comparaison de modalités d'application nationale de notre dispositif d'évaluation TGCE intercalibré, il est naturellement apparu central de tester aussi le **seuil de taille de 10 000 km²** utilisé pour typifier les TGCE d'intercalibration, afin d'étudier l'influence de son application éventuelle sur la classification d'état écologique des TGCE de France. Cela a fait l'objet du **Scénario 4** dont les résultats bruts sont récapitulés au **Tableau 8** ci-dessous.

Tableau 8 : Résultats d'évaluation diatomique au relevé selon le **Scénario 4** (application des nouvelles grilles d'évaluation sur les TGCE définis au seuil de surface > 10 000 km². Grilles 2015 sur cours d'eau plus petits)

Scénario 4				
Classe de qualité	Effectif total évalué	Pourcentage	Effectif de TGCE évalués	Pourcentage TGCE évalués
Très bon	54	8,05	38	6,91
Bon	210	31,30	164	29,82
Moyen	340	50,67	297	54,00
Médiocre	65	9,69	49	8,91
Mauvais	2	0,30	2	0,36
Total	671	100,00	550	100,00

Selon Scénario 4 : 550 TGCE (seuil 10 000 km²)

Dans ce scénario à règles d'évaluation composites, tous les cours d'eau de plus de 10 000 km² sont évalués selon les nouvelles règles d'évaluation TGCE 2018 (regroupant 550 relevés diatomiques sur cette série temporelle 2007-2013). Les cours d'eau compris entre 8 000 et 10 000 km² y sont évalués selon le dispositif de l'Arrêté 2015 dédié aux plus petits cours d'eau (soit un total de 121 relevés dans le jeu de données sélectionné).

6.1.4. Inter-comparaison et discussion des résultats des 4 scénarios

Une comparaison synthétique des résultats d'évaluation sur le jeu de données complet sélectionné (soit 671 relevés susceptibles, selon le scénario, d'être classés en TGCE) est présentée en **Figure 58** ci-dessous.

Comparaison des scénarios

n = 671 opérations sur 127 sites

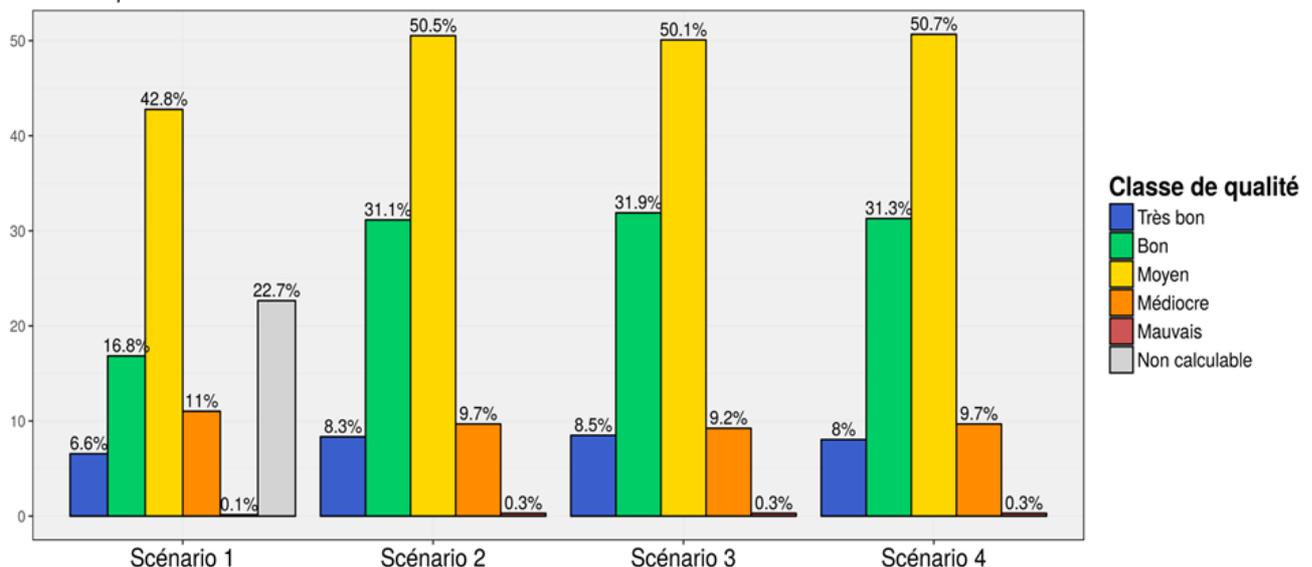


Figure 58 : Résultats comparés d'évaluation au relevé procurés par les 4 scénarios étudiés à l'échelle du jeu de données global des 671 relevés diatomiques

Il est important de rappeler que, dans 2 scénarios (**Scénario 2** et **Scénario 4**), l'évaluation sur cette assise de relevés candidats-TGCE est **composite** du fait qu'elle émerge à 2 dispositifs différents d'évaluation. En effet, la partition entre TGCE et cours d'eau plus petits conduit à l'utilisation de **2 grilles d'EQR de niveau différent** (TGCE évalués avec les nouvelles grilles TGCE 2018, cours d'eau non-TGCE évalués avec les grilles de l'Arrêté 2015).

Cette pratique d'évaluation différente selon ces 2 niveaux de taille influe mécaniquement sur le niveau d'évaluation de ces cours d'eau et sur le résultat statistique global d'évaluation produit sur ce jeu de données (671 relevés). En effet, les grilles de plus petits cours d'eau procurent une évaluation plus sévère que les grilles TGCE. Aussi, **plus il va rester de relevés et de cours d'eau évalués selon un mode « plus petits cours d'eau »**, qui mobilise des grilles plus sévères, **plus la statistique d'état a tendance à se dégrader**, le cas extrême à ce niveau étant le Scénario 1, qui n'est pas composite mais applique le dispositif exact de l'Arrêté Evaluation 2015. Sur cette base, tous les relevés ayant pu être évalués le sont selon des grilles « plus petits cours d'eau » (cf. statistique « Pourcentage sans les NA » en 4^{ème} colonne de la Figure 58, relative à la statistique d'état de tous les relevés évalués dans ce Scénario 1). C'est ce scénario initial qui donne les résultats statistiques d'état les plus sévères parmi les 4 scénarios testés (seulement 33,4 % des relevés évalués et 30,25 % de l'assise complète des relevés sélectionnés -du fait des NA, colonne grise à droite de ce scénario- dépassant le seuil du Bon Etat). Dans les 2 systèmes composites, c'est surtout la proportion de relevés évalués selon les grilles « plus petits cours d'eau » qui contribue à faire baisser globalement la statistique d'état des relevés évalués.

Les 2 autres scénarios ne font pas intervenir de système d'évaluation à 2 vitesses. Dans le **Scénario 1**, même si la ségrégation typologique entre TGCE et plus petits cours d'eau intervient, dans le système de l'Arrêté 2015, elle ne conduit pas, à une utilisation différenciée de 2 niveaux de grilles liées à ce critère de taille selon que l'on est en TGCE ou en plus petits cours d'eau (d'où sa plus grande sévérité sur les résultats globaux d'évaluation). Et dans le **Scénario 3**, tous les cours d'eau-candidats TGCE, qui ont été sélectionnés d'emblée selon le critère de surface intégrée > 8000 km², sont évalués selon des grilles spécifiques TGCE (dispositif 2018).

En complément de l'examen de cette Figure 58, qui présente le résultat comparatif global de l'application de ces 4 scénarios sur la statistique d'évaluation des relevés par classe d'état écologique, il va aussi être utile de se reporter aux Tableaux 5 à 8 pour l'étude comparative plus fine des résultats selon le scénario, notamment lorsque la partition du jeu de données entre 2 catégories de taille de cours d'eau conduit à une différence de niveau d'évaluation par l'utilisation de grilles spécifiques TGCE et d'autres grilles dédiées à des hydrosystèmes plus petits.

Ces considérations préalables sur les principes de construction des scénarios et leur incidence mécanique sur l'évaluation étant posées, étudions maintenant plus en détail les résultats comparatifs de chaque scénario.

Comme cela a été rappelé, le **Scénario 1** utilise les grilles d'évaluation de l'Arrêté du 27-07-2015, **qui n'ont pas été élaborées dans un cadre de réflexion spécifique TGCE**. En partie gauche de la Figure 58, on peut déjà remarquer qu'une part non-anecdotique du jeu de 671 relevés (représentant 22% du total) n'y est pas évaluée. Cette partie figurée en grisé correspond aux relevés prélevés sur des Très Grands Cours d'Eau Alpains (Rhin, Rhône). En effet, en l'absence de références vraies sur de tels cours d'eau dans notre territoire national et dans l'attente d'indications émanant du niveau européen sur les niveaux de référence par défaut à appliquer, l'Arrêté Evaluation 2015 ne donnait pas de cadre d'évaluation applicable pour ces types de cours d'eau (dièse dans la case correspondante). Notre récente participation nationale au nouvel exercice d'intercalibration TGCE 2015-2016 nous a donné l'occasion de bénéficier de l'expérience préalable capitalisée par le précédent X-GIG « Large Rivers » à ce niveau, et de pouvoir procéder à diverses itérations de calage du dispositif Français s'appuyant sur le niveau de référence par défaut adopté à l'issue du 1^{er} exercice X-GIG antérieur (c'est ce niveau de référence moyen qui a aussi été utilisé pour notre « Fit-in » Exercise).

Dans l'histogramme présenté en Figure 58 la statistique globale d'état écologique présentée par classe de qualité pour le Scénario 1 est impactée par ces 22 % de relevés non-évalués. Il faut donc aller se reporter au détail des exploitations statistiques données au Tableau 5 afin de comparer plus finement les résultats de ce scénario et des scénarios suivants, en gardant toujours en mémoire que la base de relevés servant à ces inter-comparaisons entre Scénario 1 et les 3 autres a une assise différente (le Scénario 1 n'évaluant aucun relevé du type national TGCA, les autres scénarios les intégrant dans leur statistique...).

Si l'on se reporte aux 2 colonnes les plus à droite de ce Tableau 5, qui établissent une statistique d'état sur les 429 relevés provenant de TGCE d'après la typo nationale, il n'est pas surprenant de constater **que ce Scénario 1 est le plus sévère des 4 scénarios inter-comparés**, avec seulement 25,6 % des relevés TGCE évalués au moins en Bon Etat (¾ des relevés n'atteignant pas le Bon Etat), la classe d'Etat Moyen contenant à elle seule plus de 60 % des relevés évalués. Ce résultat est logique et était attendu, puisque les grilles d'évaluation utilisées avaient été conçues à partir de niveaux de référence vraie calculés sur des sites réels de cours d'eau beaucoup plus petits.

En effet, les grilles de calcul d'EQR de l'Arrêté 2015 ont donc été calées, puis intercalibrées pour l'évaluation de cours d'eau plus petits qui, logiquement, sont évalués sur une base plus sévère (seuils de basculement de classes de qualité en EQR un peu plus élevés que pour les TGCE).

L'examen du reste de la Figure 58, qui est plus inter-comparable vu qu'il repose sur le même effectif de 671 relevés évalués, montre que les 3 autres scénarios d'évaluation, qui intègrent pour partie (Scénario 2 et Scénario 4) ou en totalité (Scénario 3, basé sur l'adoption en TGCE de tous les relevés obtenus sur cours d'eau de plus de 8 000 km² de BV intégré) l'utilisation des nouvelles grilles TGCE proposées dans le dispositif 2018, donnent une statistique de résultats d'évaluation très proches si l'on considère le classement d'état écologique obtenu sur le jeu de données global de 671 Opécont, tous relevés inclus dans cette statistique.

Naturellement, le Scénario 3, qui utilise les grilles de type TGCE un peu moins sévères, puisque dédiées TGCE, sur l'assortiment complet des sites sélectionnés, donne une statistique globale d'état un tout petit peu plus favorable (40,4 % des relevés évalués en Bon Etat) que le Scénario 2 (39,4%) et que le Scénario 4, le plus sévère (39,3 %). Ces écarts, qui font changer cette statistique de seulement 1 à 1,1 % de Bon Etat sur tout le jeu de données d'étude, **sont vraiment marginaux** et influenceront très peu, quel que soit le choix fait, sur le niveau de transposition et sur la validité de mise en œuvre des résultats d'intercalibration.

Si l'on veut cette fois s'intéresser de plus près à la statistique d'évaluation basée sur la **seule assise des relevés TGCE sélectionnés en fonction du scénario**, il faut inter-comparer les résultats des 2 colonnes les plus à droite (Effectif de TGCE évalués et pourcentage TGCE évalués) des Tableaux 5, 6, 7 et 8.

La tendance est exactement la même, à savoir que c'est le Scénario 3 qui donne une statistique légèrement plus favorable que les 2 autres sur l'évaluation des relevés strictement TGCE. En effet, il conduit à évaluer 40,38 % de relevés atteignant au moins le Bon Etat (sur une assise de 671 relevés), pour 37,53 % de BE sur une assise de 581 relevés dans le cas du Scénario 2 et 36,73 % de BE sur une assise de 550 relevés pour le Scénario 4, le plus sévère des 3.

Une hypothèse explicative possible à ce léger différentiel pourrait tenir au fait que nos cours d'eau Français entre 8 000 et 10 000 km² pourraient être statistiquement légèrement moins altérés que ceux passant les 10 000 km² (explication du différentiel entre Scénario 3 et Scénario 4, dont les différences découlent du seuil de taille utilisé pour la définition des TGCE), du simple fait de se trouver un peu plus en amont des bassins versants concernés.

Si cette hypothèse est envisageable, il n'est cependant pas garanti que ce résultat soit effectivement dû à cette origine. En effet, dans le monde réel, il est souvent constaté une tendance générale inverse. En effet, sauf situation préservée en amont des bassins versants, de plus petits cours d'eau peuvent subir des gradients de pollution nettement plus intenses que de plus gros systèmes aval où, même s'il est difficile d'observer de hauts niveaux de qualité abiotique et biologique du fait de l'aspect composite des bassins versants, la mixité des origines, les assez longs temps de résidence, la capacité d'auto-épuration et les aspects de dilution contribuent à modérer le degré d'impact anthropique. Dans ce cas, ce différentiel pourrait aussi être généré par le hasard de la distribution de qualité des relevés qui vont passer d'un système de taille à l'autre selon le scénario.

S'agissant cette fois des différences de résultat entre Scénario 3 et Scénario 2, l'explication est plus difficile à isoler car il est possible qu'elles soient fortement sous influence de la différence de classification de taille selon les 2 typologies utilisées, qui génère une importante zone de chevauchements entre surfaces de bassins versants et classes de taille sur base hydromorphologique (utilisation des ordres).

Dans un tel contexte, certains hydrosystèmes étant traités comme des TGCE selon le système de typification basé sur les surfaces sont considérés comme des cours d'eau plus petits dans le système basé sur les ordres et vice-versa (les 2 types de cas ayant été repérés). Cela joue sur l'assise de TGCE étant évalués en tant que tels sous un scénario ou l'autre et peut générer du flou sur la statistique d'évaluation (globale, bien sûr, par l'utilisation de grilles de sévérité différente, mais aussi sur la statistique d'évaluation spécifique TGCE, qui ne porte plus exactement sur la même assise selon le système de typification de taille employé).

La différence de nombre de relevés TGCE entre le Scénario 2 (581) et le Scénario 4 (550) illustre en partie la différence de perception basée sur un critère hydromorphologique (ordre de Strahler) et un critère de seuil de surface basé sur un descripteur à notre sens plus précis, qui entraîne une marge de chevauchement inter-types de taille nettement plus faible (du niveau de l'erreur d'évaluation ou de modélisation de surface). A ce différentiel de nombre de relevés entre ces 2 systèmes, qui est une résultante seulement partielle d'inter-croisements typologiques, il faudrait aussi rajouter des sites classés TGCE sur la base de la typo nationale et n'ayant pas été sélectionnés ici du fait de surfaces intégrées inférieures à 8 000 km², ce seuil de sélection pour notre travail dédié « mise au point d'un nouveau système d'évaluation TGCE » ayant été conditionné par le fait qu'on ne voulait pas trop s'éloigner de la vision européenne du type de taille de TGCE.

Au final, les différences de classification produites par les 3 scénarios intégrant les nouvelles règles d'évaluation TGCE (Scénarios 2 à 4) ne diffèrent pas de façon décisive sur les résultats d'évaluation au relevé produites, et les différences sont même plutôt minces entre les 3 scénarios faisant appel au nouveau système TGCE intercalibré. Parmi ces derniers, c'est le Scénario 3, qui utilise les nouvelles grilles de type TGCE sur le nombre le plus élevé de relevés, qui produit l'évaluation la moins sévère. Ainsi, 40,4% des relevés, intégrant pas mal de relevés de cours d'eau assez directement issus de zones montagneuses (Notamment Durance, Dordogne...), y seraient évalués en bon état.

Du fait du centrage assez net du jeu de données d'évaluation en-dessous du BE, avec une classe d'EM dominant l'effectif (un peu plus de 50 %), il est possible d'anticiper que la statistique d'état va subir une dégradation avec le jeu des agrégations triennales d'état, et se dégradera encore plus dans les cas des Scénarios 2 et 4 (jeu de données d'évaluation encore plus décentré vers la classe d'Etat Moyen) que pour le Scénario 3, qui aurait notre faveur.

Nous pensons aussi que, dans le cas de cette séparation entre cours d'eau plus petits et TGCE, qui conduit de façon automatique à un changement de niveau de grille d'évaluation (passage de niveaux de référence réels calculés à un niveau de référence par défaut et, dans le même temps, adoption d'une grille d'EQR plus basse pour l'évaluation), **il serait préférable d'adopter un critère clair de définition de taille basé sur les surfaces intégrées de bassin versant**, ménageant le moins possible d'inter-croisement typologique.

En effet, une hétérogénéité de détermination du type, et la zone de d'inter-croisement aléatoire de tailles de systèmes qui peut en résulter, sont de nature à provoquer des problèmes d'équité d'évaluation de cours d'eau d'importance équivalente, en fonction du bassin hydrographique dans lequel on se trouve et de la façon dont les ordres de Strahler y ont été définis.

Les **surfaces intégrées de bassin versant**, sur lesquelles sont basés 2 des scénarios d'évaluation testés (seuil > 8 000 km² pour le Scénario 3 et seuil > 10 000 km² pour le Scénario 4) représentent un descripteur légitime de taille d'hydrosystème, utilisé notamment par l'Union Européenne pour le cadrage des types d'intercalibration UE.

Le seuil de 10 000 km² est celui utilisé pour l'intercalibration Européenne. Il donne la statistique d'évaluation la plus sévère des 3 scénarios proposés, qui va encore subir une dégradation mécanique avec la pratique des moyennes triennales d'état. C'est pourquoi le Scénario 3 a été présenté dans cette note d'appui comme ayant notre préférence. En effet, il présenterait l'avantage d'**évaluer dans un système référentiel plus cohérent** les secteurs aval de plaine des grands affluents de nos fleuves (qui, objectivement, ne peuvent déjà plus être rattachés à un système de référence vraie calculée valide pour de tels cours d'eau) ainsi que le drain aval du fleuve concerné (système plus homogène et plus intercomparable).

S'agissant du risque de sous-transposition de l'intercalibration réalisée, il reste de la sphère de décision nationale de proposer une typologie d'évaluation adéquate en fonction de ses spécificités naturelles et des particularités de son réseau hydrographique, l'important étant surtout de vérifier par l'intercalibration que notre système national évalue le Bon Etat de façon non laxiste par rapport aux autres Etats-Membres et à la vision partagée du Bon Etat.

A cet effet, il faut rappeler ici que dans la plupart des exercices d'intercalibration réalisés sur plus petits cours d'eau (CB GIG, Alpine GIG et MED GIG), la France se situait la plupart du temps au moins dans le secteur de la moyenne européenne, et assez souvent dans la fourchette supérieure de biais autorisés. Dans le dernier exercice d'intercalibration TGCE réalisé, le choix final du Scénario 3-3 (cf. Figure 55) se situait dans un biais positif jugé acceptable dans le cadre de l'exercice (< 0,5), mais qui garantit une évaluation sur un niveau jugé, au sein de l'exercice, comme un peu plus sévère que la vision dominante de l'exercice auquel la France a participé.

Rappelons que ce choix de dispositif présentant des biais positifs a été justifié par des éléments particuliers à notre système d'évaluation : 1) l'IBD n'évalue pas que l'altération par les orthophosphates et l'eutrophisation, cible du présent exercice d'intercalibration, mais aussi 6 autres descripteurs d'altération anthropique ; et 2) la France présente, comme quelques autres pays, des TGCE en provenance directe de zones montagneuses, en rien comparables à de très grands cours d'eau de plaine type Danube, et leur nécessaire évaluation dans un système un peu plus sévère, mais justifié compte-tenu des différences de taille et de situation naturelle, contribue à générer des biais positifs.

Nous avons rappelé à l'époque qu'il existe un degré de souplesse assez important dans la façon dont un Etat-Membre peut typifier ses masses d'eau et adapter leur système d'évaluation en fonction des caractéristiques et particularités de son réseau hydrographique national (il y a un peu de mou dans le système, et c'est heureux...). D'autre part, le niveau de calage actuel du dispositif Français intercalibré, qu'il s'agisse du nouveau système TGCE ou des systèmes intercalibrés antérieurement sur des types européens plus petits, ménage une marge de sécurité conséquente par rapport au calage moyen de dispositif de la plupart de nos pays voisins (les biais de la France ayant pratiquement toujours été de modérément à assez nettement positifs).

Compte-tenu de ce contexte et des marges de manœuvre existantes, il a été indiqué qu'objectivement, en retenant le seuil de surface de 8 000 km², il n'y aurait pas de risque pour la France, du moins dans le domaine de l'évaluation diatomique des cours d'eau et des très grands cours d'eau, de sous-transposer l'évaluation DCE du Bon Etat et les résultats des intercalibrations réalisées.

Quoi qu'il en soit, à un moment donné, il est revenu aux tutelles de prendre la décision et de porter la responsabilité du choix final réalisé. Pour éviter tout risque de sous-transposition des intercalibrations antérieurement réalisées sur hydrosystèmes plus petits, ainsi que l'éventuel risque juridique associé, la Direction de l'Eau, en concertation avec l'AFB, a pris la décision de **retenir officiellement le seuil inférieur de surface de 10 000 km²** pour l'application des nouvelles grilles TGCE récemment intercalibrées.

La rédaction du modificatif d'Arrêté Evaluation à paraître en 2018 a donc été effectuée en ligne avec cette décision, basant l'application du nouveau dispositif TGCE sur le même critère clair de typologie que l'U.E. (surface minimale de 10 000 km²) ; les cours d'eau plus petits restant pour leur part évalués selon le dispositif d'évaluation antérieurement en vigueur (cf. contenu de l'Arrêté Evaluation du 27-07-2015).

6.2. Proposition d'évolution du contenu de l'Arrêté Evaluation 2018

Les propositions que nous avons successivement émises auprès des tutelles dans la perspective de mise en application officielle du nouveau dispositif d'évaluation peuvent se décliner en 2 phases différentes :

- Une première phase, entre Octobre 2017 et Février 2018, où nous envisageons l'édition d'un tableau d'application de l'évaluation d'état diatomique le plus simple possible, basé sur un principe de partition claire entre typologie nationale et typologie U.E. « Large Rivers » sur la base d'un seuil de surface calé à 10 000 km² sur la proposition (éventuellement amendable à 8 000 km², selon décision finale des tutelles). Ce tableau qui se limitait, comme le tableau de l'Arrêté précédent, à 5 colonnes, conduisait à l'application de la typologie nationale et des grilles antérieures d'évaluation de la qualité diatomique sur tout cours d'eau plus petit que 10 000 km², et à l'application des nouvelles grilles TGCE sur tout cours d'eau supérieur à 10 000 km².
- Une 2^{ème} phase où, la Direction de l'Eau ayant entretemps décidé de rajouter une colonne supplémentaire au tableau soumis à publication au JORF dans le but de clarifier les cas de figure possibles et d'éviter une sous-transposition éventuelle des intercalibrations antérieures, il a fallu corriger, au 4^{ème} trimestre 2018 et avec un nouvel interlocuteur MTES / Direction de l'Eau, ce tableau de conception et d'interprétation plus complexe pour les acteurs, qui avait été maquetté et finalement publié par le JORF avec des erreurs de contenu.

6.2.1. Premières propositions d'évolution du contenu de l'Arrêté Evaluation :

En vue d'introduire ce nouveau système d'évaluation des TGCE, tel qu'accepté par l'Union Européenne, dans le circuit de l'officialisation, afin qu'il puisse passer dès que possible dans le domaine de l'application réglementaire, nous avons, en date du 18-10-2017 puis du 30-01-2018, envoyé 2 évolutions de propositions de modification de l'Arrêté national Evaluation pour publication en 2018.

Ces propositions incluaient d'une part quelques amendements à la marge, relatifs à des aspects typologiques inappropriés ou décalés, par exemple pour un petit nombre de cours d'eau dont soit la classification de taille était manifestement trop décalée par rapport au critère de surface –ex : Aude, plutôt GCE- ; ou lorsque le système d'évaluation antérieur était manifestement inapproprié par rapport au bassin versant intégré de ce cours d'eau à certains sites TGCE, par exemple Vilaine aval, Meuse aval, Isère aval, Durance aval.

D'autre part, le nouveau dispositif récemment inter-calibré donne un cadre d'évaluation pour le nouveau Biotype 6, auquel se rattachent tous les TGCE dont le bassin versant provient de zones montagneuses proches (Rhône à l'amont de Lyon jusqu'à sa confluence avec la Saône ; Rhin dans la Plaine d'Alsace ; Dordogne avant l'HER 14 ; Isère aval jusqu'à sa confluence avec le Rhône ; Durance aval). Pour mémoire, jusqu'à présent, les Très Grands Cours d'Eau Alpains (Type national TGCA) ne faisaient pas encore l'objet d'une évaluation d'état écologique à partir du maillon diatomique.

Les ajouts, propositions modificatives et remarques intégrées en date du 30-01-2018 pour la modification de cet Arrêté, dont beaucoup concernent la déclinaison des grilles TGCE inter-calibrées en fonction de la typologie de nos différents très grands cours d'eau français, sont récapitulés dans l'extrait de tableaux modifiés présenté aux 2 pages suivantes (certaines décisions finales à prendre pour l'édition étant encore en réflexion à l'époque).

Extrait de fichier d'Arrêté modifié proposé par Irstea :

Phrase un peu plus haut à modifier :

Le tableau 24 ci-dessous indique les valeurs inférieures des limites de classe, en EQR, pour l'indice biologique diatomées. Le tableau 25, ci-dessous indique, par type de cours d'eau, la valeur de référence **calculée ou retenue par type** et la valeur minimale sous la forme suivante : a-b (a = valeur de référence **vraie ou par défaut** du type, b = valeur minimale du type).

Etc...

Tableau 25 : valeurs de référence (vraie ou par défaut) et valeurs minimales, par type de cours d'eau, pour l'IBD₂₀₀₇

Valeur de référence et valeur minimale par type (IBD ₂₀₀₇)		Catégories de taille de cours d'eau				
		Cas général, cours d'eau exogène de l'HER de niveau 1 indiquée ou HER de niveau 2	Très Grands (> 10 000 km ²)	Grands et Très Grands < 10 000 km ²	Moyens	Petits
Hydroécocorégions de niveau 1						
20	DEPOTS ARGILO SABLEUX	Cas général			18,1-1	
		Exogène de l'HER 9			18,1-1	
		Exogène de l'HER 21			19-5	
21	MASSIF CENTRAL NORD	Cas général			19-5	
3	MASSIF CENTRAL SUD	Cas général			19-5	
		Exogène de l'HER 19			(#) 18,1-1	
		Exogène de l'HER 8			(#) 19-5	
		Exogène de l'HER 19 ou 8		18,1-1		
17	DEPRESSIONS SEDIMENTAIRES	Cas général			17,4-1	
		Exogène de l'HER 3 ou 21	(#) 18,1-1		19-5	
15	PLAINE SAONE	Exogène de l'HER 3 ou 21			19-5	
		Exogène de l'HER 5			20-5	
		Cas général	(#) 18,1-1		18,1-1	
		Exogène de l'HER 4	(#) 18,1-1			
5	JURA / PRE-ALPES DU NORD	Cas général			20-5	
		Exogène de l'HER 2	(#) 19,1-1		20-5	
TTGA	FLEUVES ALPINS	Rhin, Rhône jusqu'à Lyon	(#) 19,1-1			
TTGA	FLEUVES ALPINS	Rhône de l'aval confluence Saône à Lyon jusqu'à l'exutoire	(#) 18,1-1			
2	ALPES INTERNES	Cas général			20-5	
7	PRE-ALPES DU SUD	Cas général			20-5	
		Exogène de l'HER 2	(#) 19,1-1		20-5	
6	MEDITERRANEE	Exogène de l'HER 2 ou 7			20-5	
		Exogène de l'HER 7			20-5	
		Exogène de l'HER 8			19-5	
		Exogène de l'HER 1	(#) 18,1-1		20-5	

Commentaire [DF1]: Pour toutes les cases TGCE de ce tableau, c'est un niveau de référence par défaut qui a été appliqué selon le type, calé en fonction du niveau moyen de référence par défaut qui est ressorti collectivement de l'exercice d'intercalibration TGCE

Commentaire [DF2]: Cases TGCE existantes dans l'ancien Arrêté mais à virer et à griser du fait qu'on se réfère au seuil inférieur de TGCE à 10 000 km² (cours d'eau trop petits repassant en GCE juste à droite).

		Cas général		18,1-1
8	CEVENNES	Cas général		19-5
		A-her2 n°70		19-5
16	CORSE	A-her2 n°22		19-5
		B-her2 n°88		
19	GRANDS CAUSSES	Cas général		18,1-1
		Exogène de l'HER 8		19-5
11	CAUSSES AQUITAINS	Cas général		18,1-1
		Exogène de l'HER 3 et/ou 21	(#) 19,1-1	19-5
14	COTEAUX AQUITAINS	Exogène des HER 3, 8, 11 ou 19	(#) 18,1-1	18,1-1
		Exogène de l'HER 3 ou 8		18,1-1
		Cas général		18,1-1
		Exogène de l'HER 1	(#) 18,1-1	20-5
13	LANDES	Cas général		18,4-5
1	PYRENEES	Cas général		20-5
12	ARMORICAIN	A-Centre-Sud	(#) 17,4-1	17,4-1
		B-Ouest-Nord Est		
TTGL	LA LOIRE	Cas général	(#) 18,1-1	
9	TABLES CALCAIRES	A-her2 n°57		18,1-1
		Cas général		18,1-1
		Exogène de l'HER 10		18,1-1
		Exogène de l'HER 21	(#) 18,1-1	19-5
10	COTES CALCAIRES EST	Exogène de l'HER 21		
		Cas général	(#) 18,1-1	18,1-1
		Exogène de l'HER 4		19-5
4	VOSGES	Cas général		19-5
22	ARDENNES	Exogène de l'HER 10	(#) 18,1-1	
		Cas général		17,4-1
18	ALSACE	Cas général		18,1-1
		Exogène de l'HER 4		19-5

Commentaire [DF3]: Case fusionnée à séparer (pour la case TGCE, niveau de référence par défaut au lieu de référence vraie)

Commentaire [DF4]: Le tableau contenait une erreur par rapport au tableau récapitulatif le dispositif d'évaluation diatomique, transmis en date du 19-10-2017. Ce cours d'eau est la Meuse (seul TGCE de cette HER), qui est exogène de l'HER 10 (carbonatée). Son EQR doit bien être calculé selon cette gamme de variation

(#) : niveau de référence par défaut (absence de référence vraie). En grisé : type inexistant

a-b : a = valeur de référence (vraie ou par défaut) ; b = valeur minimale

Les valeurs de l'IBD₂₀₀₇ figurant dans ce tableau ont pris en compte :

- la décision de la commission du 20 septembre 2013 relative à l'inter-étalonnage pour les cours d'eau plus petits que les TGCE (< 10 000 km²)
- la décision de la commission du 22 mars 2017 relative à l'inter-étalonnage pour les TGCE (> 10 000 km²)

¹ Pour les très-grands-cours-de-bassin-versant supérieur à 10 000, il faut utiliser les valeurs de référence et minimale suivantes : 18,1-1

² Pour les très-grands-cours-de-bassin-versant supérieur à 10 000 km², il faut utiliser les valeurs de référence et minimale suivantes : 19,1-1

³ Pour les très-grands-cours-de-bassin-versant supérieur à 10 000 km², il faut utiliser les valeurs de référence et minimale suivantes : 17,4-1

Dans les 2 mois suivant cette proposition des spécialistes-maillon Irstea, réalisée dans le cadre de l'Appui à la Décision Publique et en vue de la publication prochaine de l'Arrêté au JORF, 2 décisions ont été prises par les tutelles :

- Adoption du seuil d'intercalibration Européenne de 10 000 km² pour l'application du nouveau dispositif d'évaluation national TGCE (en balance avec un autre scénario d'application mis en comparaison dans le cadre du dernier volet de cette étude nationale, qui se basait sur un seuil de 8 000 km²),

- Modification de la rédaction de ce tableau d'Arrêté en y rajoutant une 6^{ème} colonne.

Les objectifs ayant sous-tendu ces décisions des tutelles étaient de mieux décrire la partition entre typologie nationale antérieure et typologie européenne d'une part, d'éviter les risques de sous-transposition des intercalibrations antérieurement réalisées d'autre part.

6.2.2. Nouvelle phase de correction de l'Arrêté Evaluation du 27 Juillet 2018 :

Comme indiqué précédemment, du fait que son propre format de maquettage et d'édition ne correspondait pas à celui dans lequel la Direction de l'Eau nous a demandé de travailler sur le projet de modificatif d'Arrêté, le service de maquettage et d'édition du JORF a introduit quelques erreurs pendant le re-maquettage du nouveau tableau d'évaluation diatomique à partir de l'IBD₂₀₀₇.

Ces erreurs d'édition ayant été détectées postérieurement à la publication au J.O. (28 Août 2018), il a été nécessaire de corriger et valider à nouveau le tableau de dispositif diatomique (entre Octobre 2018 et Janvier 2019), d'une part en vue de pouvoir publier un correctif à cet Arrêté, et d'autre part afin de re-valider avec l'AFB et le nouveau correspondant du MTES certaines modifications apportées depuis le dispositif antérieur, pour intégration adéquate dans le nouveau Guide REEE et dans l'outil de calcul de l'Etat Ecologique du S3E.

Le document présenté dans les 2 pages suivantes correspond aux tableaux d'évaluation diatomique des cours d'eau contenus dans la maquette pré-finale du **nouveau Guide national REEE 2019** (version amendée diffusée largement par la Direction de l'Eau au moyen d'un mailing circulaire en date du 14-01-2019), après intégration des corrections d'erreurs et d'autres petites modifications apportées à la présentation de ces tableaux. Leur contenu reprend à l'identique celui du modificatif d'Arrêté Evaluation à paraître.

Irstea a été sollicité à plusieurs reprises entre le quatrième trimestre 2018 et la mi-janvier 2019 et a largement contribué à la vérification, à la correction et à la validation de ces tableaux destinés à l'application réglementaire de l'Evaluation.

Comme on peut le remarquer, les tableaux du Guide REEE, comme les tableaux de l'Arrêté Evaluation publiés à l'été 2018, s'appuient désormais sur un **dispositif d'évaluation basé sur 6 colonnes** en rapport avec la typologie de taille des cours d'eau.

Les 5 colonnes de droite reprennent le dispositif d'évaluation antérieur, qui utilise la typologie abiotique nationale des cours d'eau telle que mise en place au début des années 2000 par le LHQ (Equipe Irstea basée à Lyon).

Pour évaluer ces 5 colonnes, le dispositif utilise la 1^{ère} grille d'EQR donnée dans le 1^{er} tableau, dédiée à l'évaluation de **tous types de cours d'eau**, à l'exception des TGCE > à 10 000 km². Les cours d'eau qui étaient rattachés à un type TGCE dans le cadre de la typologie nationale antérieure, mais dont la surface au site de surveillance est inférieure à 10 000 km, restent évalués selon le dispositif antérieurement en vigueur (Arrêté du 27 Juillet 2015). Cette pratique évite les risques de sous-transposition des intercalibrations réalisées antérieurement. En effet, la grille d'EQR destinés aux cours d'eau plus petits est plus sévère (valeurs-seuils inter-classes de qualité plus hauts) que la grille récemment intercalibrée au niveau de l'UE pour les TGCE.

Le nouveau dispositif d'évaluation récemment intercalibré vient donc se greffer au dispositif antérieur au moyen de la 6^{ème} colonne à gauche du tableau. Il concerne les TGCE **strictement supérieurs à 10 000 km²** de surface intégrée de bassin versant, ce qui correspond au critère de définition typologique utilisé par l'Union Européenne. Il est donc important de déterminer, au long d'un linéaire d'hydrosystème fluvial donné, à partir de quelle station ce seuil de surface intégrée est atteint.

Pour les cours d'eau relevant de cette 6^{ème} colonne, c'est la 2^{ème} grille du 1^{er} tableau de valeurs-seuils en EQR, aux seuils un peu abaissés par rapport à la 1^{ère}, qui est utilisée pour une évaluation DCE-compatible.

Quelques problèmes particuliers rencontrés ont nécessité une réflexion spécifique, en vue d'adapter les conditions de mise en œuvre et de permettre une évaluation plus judicieuse de certains de nos hydrosystèmes.

Les adaptations ont notamment concerné les cas suivants :

- Le Rhône, à l'amont de sa pénétration dans l'HER 6 (domaine méditerranéen carbonaté), qui intervient dans le secteur de sa confluence avec l'Isère, appartenait jusqu'à présent, au sens de la typologie abiotique nationale, au type unique TTGA (Très Très Grands Alpains), qui regroupait aussi le Rhin sur tout son parcours Français. Or le Rhône présente une assez forte différence typologique, hydrochimique et de flores diatomiques entre son secteur Jurassien à l'amont de Lyon et l'aval de sa confluence avec la Saône, rivière de plaine dont les eaux sont beaucoup plus minéralisées et viennent se mélanger au Rhône à Lyon.

Extrait « Réactualisation du Guide National REEE », version actualisée du 14-01-2019 (2 pages) :

**Indice pour le phytobenthos : indice biologique diatomées
(code Sandre : 5856)**

La note en EQR se calcule comme suit :

Note en EQR = (note observée – note minimale du type) / (note de référence du type – note minimale du type)

Valeurs inférieures des limites des classes d'état, exprimées en EQR, par type de cours d'eau pour l'IBD

Valeurs inférieures des limites de classe de l'IBD	Limites des classes d'état IBD en EQR			
	Très bon / Bon	Bon / Moyen	Moyen / Médiocre	Médiocre / Mauvais
Tous types de cours d'eau sauf TGCE > 10 000 km ² de bassin versant	0,94	0,78	0,55	0,3
Très grands cours d'eau ≥ 10 000 km ² de bassin versant (**)	0,92	0,76	0,52	0,26

Les valeurs d'EQR de l'IBD figurant dans ce tableau ont pris en compte la décision de la commission du 12 février 2018 relative à l'inter-étalonnage.

Valeurs de référence et valeurs minimales, par type de cours d'eau, pour l'IBD

Valeur de référence et valeur minimale par type (IBD)		Catégories de taille de cours d'eau						
Hydroécorégions de niveau 1		Cas général, cours d'eau exogène de l'HER de niveau 1 indiquée ou HER de niveau 2	Très grands (**) ≥ 10 000 km ²	Très grands (*) < 10 000 km ²	Grands	Moyens	Petits	Très Petits
20	DEPOTS ARGILO SABLEUX	Cas général			18,1-1			
		Exogène de l'HER 9			18,1-1			
		Exogène de l'HER 21			19-5			
21	MASSIF CENTRAL NORD	Cas général			19-5			
3	MASSIF CENTRAL SUD	Cas général			19-5			
		Exogène de l'HER 19				18,1-1		
		Exogène de l'HER 8				19-5		
17	DEPRESSIONS SEDIMENTAIRES	Exogène de l'HER 19 ou 8			18,1-1			
		Cas général			17,4-1			
		Exogène de l'HER 3 ou 21	18,1-1	18,1-1	19-5			
15	PLAINE SAONE	Exogène de l'HER 3 ou 21				19-5		
		Exogène de l'HER 5			20-5			
		Cas général	18,1-1	18,1-1	18,1-1			
		Exogène de l'HER 4	18,1-1	18,1-1				
5	JURA / PRE-ALPES DU NORD	Cas général			20-5			
		Exogène de l'HER 2	19,1-1	20-5	20-5			
TTGA	FLEUVES ALPINS (***)	Rhin, Rhône jusqu'à Lyon	19,1-1					
TTGA	FLEUVES ALPINS (****)	Rhône de l'aval confluence Saône à Lyon jusqu'à l'exutoire	18,1-1					
2	ALPES INTERNES	Cas général			20-5			
7	PRE-ALPES DU SUD	Cas général			20-5			
		Exogène de l'HER 2	19,1-1	20-5	20-5			
6	MEDITERRANEE	Exogène de l'HER 2 ou 7			20-5			
		Exogène de l'HER 7			20-5			

		Exogène de l'HER 8		18,1-1	19-5		
		Exogène de l'HER 1			20-5		
		Cas général			18,1-1		
8	CEVENNES	Cas général			19-5		
		A-her2 n°70			19-5		
16	CORSE	A-her2 n°22			19-5		
		B-her2 n°88					
19	GRANDS CAUSSES	Cas général			18,1-1		
		Exogène de l'HER 8			19-5		
11	CAUSSES AQUITAINS	Cas général			18,1-1		
		Exogène de l'HER 3 et/ou 21	19,1-1	19-5	19-5		
14	COTEAUX AQUITAINS	Exogène des HER 3, 8, 11 ou 19	18,1-1	18,1-1	18,1-1		
		Exogène de l'HER 3 ou 8			18,1-1		
		Cas général			18,1-1		
		Exogène de l'HER 1	18,1-1	18,1-1	20-5		
13	LANDES	Cas général			18,4-5		
1	PYRENEES	Cas général			20-5		
12	ARMORICAIN	A-Centre-Sud		17,4-1	17,4-1		
		B-Ouest-Nord Est					
TTGL	LA LOIRE	Cas général	18,1-1	18,1-1			
9	TABLES CALCAIRES	A-her2 n°57			18,1-1		
		Cas général	18,1-1		18,1-1		
		Exogène de l'HER 10			18,1-1		
		Exogène de l'HER 21	18,1-1	18,1-1	19-5		
10	COTES CALCAIRES EST	Exogène de l'HER 21			18,1-1		
		Cas général	18,1-1	18,1-1	19-5		
		Exogène de l'HER 4					
4	VOSGES	Cas général			19-5		
22	ARDENNES	Exogène de l'HER 10	18,1-1	18,1-1			
		Cas général			17,4-1		
18	ALSACE	Cas général			18,1-1		
		Exogène de l'HER 4			19-5		

(*) : Cours d'eau classés TGCE selon la typologie nationale, mais dont la surface intégrée de bassin versant n'atteint pas 10 000 km² au site d'observation
(**) : Cours d'eau classés TGCE selon la typologie européenne du GIG Large Rivers (tous cours d'eau dont la surface intégrée de bassin versant atteint ou dépasse 10 000 km² au site d'observation
(***) : Le type TTGA de la typologie nationale, pour les masses d'eau se trouvant dans le Rhin, ou dans le Rhône à l'amont de Lyon, pourra être désigné TTGA1 »
(****) : Le type TTGA de la typologie nationale, pour les masses d'eau se trouvant dans le Rhône à l'aval de sa confluence avec la Saône à Lyon, pourra être désigné « TTGA2 »
En grisé : type inexistant
a-b : a = valeur de référence ; b = valeur minimale
Les valeurs de l'IBD figurant dans ce tableau ont pris en compte la décision de la commission du 12 février 2018 relative à l'inter-étalonnage pour les cours d'eau.

Donc, afin d'évaluer judicieusement l'état écologique du Rhône sur ce tronçon « amont », dont la typologie diatomique et le niveau de référence par défaut sont différents à partir de sa confluence avec la Saône, il a été nécessaire de subdiviser ce type national TTGA, qui n'est pas suffisamment homogène sur le plan abiotique et biologique, en 2 entités typologiques différentes, qui doivent être évaluées selon une grille différente (TGCE de Type 6 *versus* TGCE de Type 4).

- Evaluation de certains cours d'eau à ré-adapter :

- en fonction de leur origine hydrochimique amont principale (cas de la Meuse, s'écoulant dans l'HER 22 peu minéralisée « Ardennes », mais dont le bassin versant amont provient de façon très dominante d'une HER carbonatée) ;
- du fait de critères de surface connus jusqu'à présent de façon imprécise (cas de la Meuse à Givet, qui intègre une partie de BV belge auparavant non-comptabilisé, cas de la Vilaine à St Dolay qui dépasserait un peu le seuil de 10 000 km² etc...).

La fin de l'Appui à Décision Publique réalisé par Irstea sur ce chantier a consisté en :

- la rédaction de la Note d'expertise et d'appui à décision Publique présentant les 4 scénarios différents de mise en application du nouveau dispositif d'évaluation des TGCE et de leurs résultats comparés (2 versions successives dont la version finale du 26-02-2018 intégrant notamment la correction d'une erreur de citation de scénario signalée par Yorick Reyjol, qui nuisait à sa bonne compréhension),
- la tenue, dans les 2 semaines qui ont suivi, de 2 réunions triangulaires de concertation (Irstea, Direction de l'Eau, AFB) en visioconférence, afin de débattre sur les possibilités de choix en vue de la mise en application des modalités pratiques concrètes de déclinaison du dispositif d'évaluation TGCE récemment intercalibré, leurs avantages, leurs inconvénients.

Entre fin Février et la 1^{ère} quinzaine de Mars 2018, le choix final de seuillage du type TGCE à 10 000 km² a été décidé par la Direction de l'Eau, qui a consolidé en conséquence le projet de nouvel Arrêté Evaluation, puis l'a mis dans le circuit des signatures réglementaires.

Des erreurs de parution dans certains tableaux, dont le tableau d'évaluation diatomique des cours d'eau, ont rendu nécessaire la correction du contenu de cet Arrêté, ce qui a entraîné à nouveau plusieurs navette du document afin de rectifier les aspects inexacts et d'en valider le contenu (entre Octobre-Novembre 2018 et Janvier 2019).

7. CONCLUSIONS

Le présent rapport final récapitule les éléments et résultats obtenus dans le cadre d'une étude nationale visant à la mise en place d'un **nouveau dispositif d'évaluation des Très Grands Cours d'Eau de France**. Cette action confiée à Irstea a été soutenue dans le cadre d'une fiche-action prenant place dans le dispositif partenarial AFB-Irstea (Fiche-Action N° 23 « Amélioration des méthodes végétales en cours d'eau » (2016-2018)).

Dans ce cadre, en s'appuyant sur le gisement maximum de données pré-existantes mobilisables (645 relevés diatomiques concernant des cours d'eau de plus de 8 000 km² et 521 relevés de plus de 10 000 km², un nouveau dispositif d'évaluation des TGCE de France a pu être proposé et calé à partir du travail réalisé en 2015 et 2016, en tirant le meilleur parti de la participation nationale à un exercice d'intercalibration « Large Rivers » organisé au niveau Européen et de la vision partagée du Bon Etat des TGCE qui s'en dégageait. Cette participation a permis de régler directement dans une vision commune une difficulté méthodologique, à savoir le calage adéquat de niveaux de référence par défaut sur de tels cours d'eau, en l'absence de situations de référence vraie calculables sur notre territoire national (comme, du reste, de la plupart de nos pays voisins...).

L'exercice européen d'intercalibration TGCE a été visé OK par l'UE (information orale officielle en Octobre 2017, puis Décision écrite officielle UE en Février 2018), attestant de la « DCE-compatibilité » du nouveau dispositif Français proposé et intercalibré.

Le nouveau dispositif d'évaluation des TGCE de France à appliquer, qui correspond au **Scénario 3-3 du 16-06-2016** finalement présenté à l'intercalibration, suite à la décision prise de concert avec la Direction de l'Eau et l'AFB, est celui présenté en **Figure 56** et rappelé à la rubrique « Synthèse pour l'opérationnel » en début de rapport.

La rédaction d'un Arrêté Evaluation modificatif de l'Arrêté du 27-07-2015, à appliquer pour la 2^{ème} moitié du PGME en cours (Plan de Gestion des Masses d'Eau 2016-2021) a été réalisée entre Octobre 2017 et Février 2018. Cet Arrêté, signé en date du 27 Juillet 2018 et paru au JORF en date du 28-08-2018, a permis d'introduire ce nouveau dispositif d'évaluation des TGCE dans la réglementation Française et de lui donner un cadre officiel d'application.

Dans le cadre de l'étude nationale « Evaluation des TGCE », un travail portant sur les modalités pratiques de mise en œuvre du nouveau dispositif a permis de tester l'influence de 4 scénarios d'application sur les résultats d'évaluation de l'état écologique des TGCE (671 relevés diatomiques prélevés sur 127 sites différents candidats au statut de TGCE, selon seuil de taille mis en œuvre).

Le **Scénario 1** (= scénario contemporain), donné à simple titre indicatif, correspondait à l'application du dispositif d'évaluation actuellement en vigueur avant introduction du nouveau dispositif intercalibré (donc non-valable pour l'application future).

Le **Scénario 2** permettait de simuler l'introduction des nouvelles grilles en fonction de la classification TGCE ou non-TGCE basée sur la **typologie hydromorphologique nationale** (classification selon les Ordres de Strahler).

Deux autres nouveaux scénarios se sont appuyés sur un seuil typologique basé sur la **surface intégrée de bassin versant au site d'observation** (seuil > 8 000 km² pour le **Scénario 3**, concernant 671 relevés diatomiques candidats-TGCE) et seuil > 10 000 km² pour le **Scénario 4**, considérant 550 relevés diatomiques TGCE) afin de définir le rattachement ou non à la catégorie TGCE (cf. pratique de définition des types européens d'intercalibration). L'affectation d'un site donné à la catégorie « TGCE » ou « plus petit que TGCE » influe directement sur le niveau d'évaluation du site du fait de la différence de sévérité de la grille applicable.

Logiquement, c'est le Scénario 3 qui procurait des résultats d'évaluation un peu moins sévères, utilisant des grilles TGCE, donc moins sévères, sur la totalité de la sélection de relevés de plus de 8 000 km² de BV (soit 671 relevés). Le Scénario 2 donnait des résultats intermédiaires et le Scénario 4 (seuil TGCE : 10 000 km²) se révélait le plus sévère. Toutefois, les différences de statistique d'évaluation obtenues par les 3 derniers scénarios étaient **assez marginales** et, de l'avis d'Irstea, aucun de ces 3 dispositifs ne faisait encourir de risque de recours européen au motif de sous-transposition des intercalibrations réalisées.

Pour la typification de taille des TGCE, quelle que soit la taille-seuil choisie au final, nous avons recommandé d'adopter un descripteur typologique basée sur la **surface intégrée de BV**, descripteur plus objectif générant le moins de chevauchements croisés inter-types (l'erreur de mesure ou de modélisation étant en général contenue dans une fourchette d'1 à 2% maximum et, argument important, descripteur reconnu et utilisé par l'UE pour réaliser les tris typologiques préalables à la réalisation d'exercices d'intercalibration homogènes).

Par comparaison, la typologie nationale basée sur un descripteur hydromorphologique (les Ordres de Strahler) génère des chevauchements inter-catégories de tailles pouvant passer de 1 à 4 entre les plus petits TGCE hors outlier (Aude...) et les plus grands GCE (la Maine), du moins dans notre contexte national ou chaque bassin a dû produire sa propre couche hydro, avec des méthodologies probablement trop peu homogénéisées dès le départ et l'introduction de la dose d'effet-opérateur qui va avec.

Le **Scénario 3** d'application des résultats d'intercalibration (seuil TGCE > 8 000 km²), qui présentait plusieurs avantages, aurait eu notre faveur et a fait l'objet d'une recommandation de notre part :

- A l'instar du Scénario 4, la ségrégation de taille entre GCE et TGCE est générée sur la base d'un **critère typologique clair** (la surface intégrée de BV au site d'observation), ménageant un plus faible degré d'erreur de classification et de chevauchement de types de taille que la typologie nationale pré-existante. Cela doit conduire à moins d'hétérogénéité aléatoire d'évaluation d'Etat, au sein d'un même bassin hydrographique ainsi qu'en inter-bassin, pour une simple raison d'affectation décalée de taille de système dont va découler l'application d'une grille d'évaluation plus ou moins sévère pour cette seule raison. La typologie de taille basée sur la surface de BV génère donc nettement moins d'imprécision et de chevauchements typologiques aléatoires et permet ainsi de garantir une meilleure équité d'évaluation, notamment en inter-bassins.

- Il s'adressait à de **gros hydrosystèmes situés bien à l'aval** dans le continuum fluvial national, qui ont souvent déjà parcouru un linéaire important dans des secteurs de plaine et présentent des cortèges biologiques typiques de ces conditions, vis-à-vis desquels la France n'a plus aucune situation de référence vraie existante. Ces situations se trouvent déjà largement dans la même logique d'absence de références valides que pour des TGCE un peu plus gros ;

- Il présentait en outre l'**avantage d'évaluer dans un système référentiel cohérent les secteurs aval de plaine des grands affluents de nos fleuves**, qui n'ont plus de système de référence vraie valide, et le **drain aval du fleuve concerné** (système d'évaluation plus homogène et plus inter-comparable dans une logique de continuum fluvial du même ordre de taille).

- La différence de statistique d'évaluation entre ce seuil de surface à 8 000 km² et le seuil de 10 000 km² (seuil utilisé pour cadrer les exercices d'intercalibration Européenne Large Rivers, qui a servi de base au Scénario 4) n'est pas très conséquente, mais procure un état statistiquement un peu plus dégradé avec le Scénario 4 qu'avec le Scénario 3. L'examen de notre situation nationale et de nos biais positifs (*i.e.* dispositif Français paraissant plus sévère), par rapport à la moyenne de l'exercice d'intercalibration TGCE auquel la France a participé, ainsi que des intercalibrations précédentes, permettait d'assurer l'absence de risque de sous-transposition de la DCE et des résultats d'intercalibration. En effet, la France présentait déjà des biais positifs sur la plupart des cours d'eau plus petits que TGCE intercalibrés, selon le cas, dans le CB-GIG (Groupe Géographique d'Intercalibration « Central-Baltique), dans l'Alpine GIG, dans le Mediterranean GIG.

- Enfin, un dernier élément qui nous semblait à intégrer et à anticiper, compte-tenu de l'équilibre interne du jeu de données d'évaluation dont le barycentre se situe déjà largement dans la classe d'Etat Moyen, était que la statistique d'évaluation des sites sur une base de moyenne triennale va se dégrader par rapport à la statistique d'évaluation au relevé individuel telle que présentée.

Quoi qu'il en soit, dans le printemps 2018, les tutelles ont pris la décision finale de retenir le **Scénario 4** de mise en œuvre du dispositif d'évaluation (**seuil typologique de la catégorie TGCE commençant à partir de 10 000 km²**), qui a donc servi de base pour la déclinaison du nouvel Arrêté Evaluation 2018.

Ce choix en conformité exacte avec la typologie d'intercalibration UE, évite toute éventualité de recours juridique au niveau européen ; l'inconvénient étant cependant qu'il va procurer, dans le cadre du rapportage Européen d'état écologique, une évaluation de l'état écologique de nos TGCE nationaux d'un niveau qui paraissant un peu plus sévère par rapport à tous nos Pays-Membres voisins.

8. SIGLES & ABBREVIATIONS

[Sigles ou abréviations de termes utilisés dans le document. Chaque sigle ou abréviation est décrit sur une ligne, avec son libellé et la déclinaison complète de sa signification. Les sigles et abréviations sont mélangés, mais sont tous classés par ordre alphabétique.]

Alk. : Alcalinity (= alcalinité de l'eau)

DCE (=WFD) : Directive Communautaire sur l'Eau (= Water Framework Directive) : cf **Ref 1**.

d-ICM : diatom Intercalibration Common Metric (métrique commune d'intercalibration diatomique)

ECOSTAT : Groupe de Travail Européen chargé d'approches biostatistiques et méthodologiques appliquées aux milieux aquatiques (animé par le JRC : Join Research Center de l'U.E.)

EQB (= BQE) : Elément de Qualité Biologique (= Biological Quality Element)

EQR : Equivalent-Quality Ratio (= équivalent de qualité écologique)

Fit-in exercise : Exercice d'adaptation (exercice d'intercalibration *post-hoc*, qui vise à évaluer la convergence d'un nouveau système national d'évaluation à un exercice d'intercalibration antérieur dont les résultats ont été validés par l'U.E.)

GCE : Grands Cours d'Eau (typologie nationale)

GIG : Geographical Intercalibration Group (= Groupe Géographique d'Intercalibration)

IBD₂₀₀₇ : Indice Biologique Diatomées, version 2007

HER : Hydro-Eco-Région (région naturelle définie selon des critères abiotiques)

ICM : Intercalibration Common Metric (= métrique commune d'intercalibration)

ICM-bm : Intercalibration Common Metric – benchmark (étalon commun d'intercalibration). Dans le cadre de notre exercice d'intercalibration, ce vocable sous-entend qu'il s'agit d'un étalon commun dont la réponse a été modélisée dans le cadre du gradient abiotique d'altération de l'exercice d'intercalibration précédent

JORF : Journal Officiel de la République Française

LDC : Least Disturbed Condition (= niveau de référence par défaut)

« **LR** » **X-GIG** : Groupe Géographique d'Intercalibration regroupé « Large Rivers »

MNT : Modèle Numérique de Terrain

MS : Member State (Pays-Membre de l'Union Européenne)

Opécont-Chimie, Opécont-Bio : Opération de Contrôle (*i.e.* de prélèvement et/ou de mesure chimique, de prélèvement biologique)

PGME : Plan de Gestion des Masses d'Eau

SD : Seuil de Détection analytique

SDAGE : Schéma Directeur d'Aménagement et de Gestion des Eaux

SEEE ou S3E : Système d'Evaluation de l'Etat des Eaux

SIE : Système d'Information sur l'Eau

SIG : Système d'Information Géographique

SOM : Self-Organizing Map (Carte auto-organisante, réseau neuronal non-supervisé)

SQ : Seuil de Quantification (valeur minimale quantifiable avec fiabilité dans le cadre de la méthode analytique utilisée)

TAC : Titre Alcalimétrique Complet (unité Française)

TTGA : Très Très Grands (cours d'Eau) Alpains (type national)

TGCE : Très Grands Cours d'Eau (= Large Rivers dans le cadre européen)

9. BIBLIOGRAPHIE

REF. 1 : DCE, Directive Cadre sur l'Eau du 23 Octobre 2000 (JOCE L 327 du 20 /12/2000). Directive 2000/60/EC du Parlement Européen et du Conseil établissant le cadre de l'action Communautaire dans le domaine de l'eau. http://ec.europa.eu/environment/water/water-framework/index_en.html

REF. 2 : Arrêté du 27 juillet 2015 modifiant l'arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement. Journal Officiel de la République Française n°0198 du 28 août 2015.

REF. 3 : Schöll, F., Birk, S., Böhmer, J. et al, 2012. XGIG Large River Intercalibration Exercise / WFD Intercalibration - Phase 2: Milestone 6 Final Report (06 September 2012). *European Commission / Joint Research Centre Eds.* 73 p.

REF. 4 : Guéguen, J., Rosebery, J., Coste, M., Delmas, F. - 2015. French Expertise Report for the LR-GIG (2015-12-15): River typology, reference situation and ecological status classification: French method based on phytobenthos (diatom communities). 14 p.

REF. 5 : Guéguen, J., Rosebery, J., Coste, M., Delmas, F. - 2015. French Expertise Report (2015-12-21) for the LR-GIG: IBD 2007: Context, included abiotic variables, pressure-impact relationships. 16 p.

REF. 6 : Coste, M. in CEMAGREF (1982). Etude des méthodes biologiques d'appréciation quantitative de la qualité des eaux (dont IPS). Rapport QE Lyon-Bassin Rhône-Méditerranée-Corse.

REF. 7 : Coste, M., Boutry, S., Tison Rosebery, J., Delmas, F. (2009). Improvements of the Biological Index (BDI): description and efficiency of the new version (BDI-2006). *Ecological Indicators*, vol. 9, n° 4, p. 621-650.

REF. 8 : van Dam, H., Mertens, A., Sinkeldam, J. (1994). A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Aquatic Ecology* 28, 117–133.

REF. 9 : Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 1980, Vol. 37, N° 1 : p. 130-137.

REF. 10 : Wasson, J.G., Chandesris, A., Pella, H., Blanc, L. (2004). Les hydro-écorégions : une approche fonctionnelle de la typologie des rivières pour la Directive cadre européenne sur l'eau. *Ingénieries - E A T, IRSTEA éditions*, 2004, p. 3-10.

REF. 10 : Rosebery, J., Bottin, M., Coste, M., Delmas, F. 2011. Typologie des flores diatomiques de référence et seuils d'état écologique à partir des notes d'IBD₂₀₀₇. Rapport final d'Etude Irstea-ONEMA. IRSTEA éditions, 2011, 17 p.

REF. 11 : European Commission / Common Implementation Strategy (CIS) / Working Group 2.3. (2003). REFCOND Guidance Nr.10: "Rivers and Lakes – Typology, Reference Conditions and Classification Systems". European Commission, Technical Report-2003, 87p.

REF. 12 : Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L. et al. (2009). A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. *Hydrobiologia* 621:169–82.

REF. 13 : Rott E., P. Pfister, H. Van Dam, E. Pipp, K. Pall, N. Binder, K. Ortler (1999). Indikationslisten für Aufwuchsalgen. Teil 2: Trophieindikation sowie geochemische Präferenz, taxonomische und toxikologische Anmerkungen. Bundesministerium für Land- und Forstwirtschaft, Wien. Publ. Wasserwirtschaftskataster, BMFLF: 1–248.

REF. 14 : European Commission (2011). Guidance document on the intercalibration process. Guidance Document Nr. 14: Implementation strategy for the Water Framework Directive (2000/60/EC). European Commission, Technical report-2011-045.

REF. 15 : European Commission (2015). Guidance document on the intercalibration process / Guidance Document Nr. 30: Procedure to fit new or updated classification methods to the results of a completed intercalibration exercise. European Commission, Technical report-2015-085.

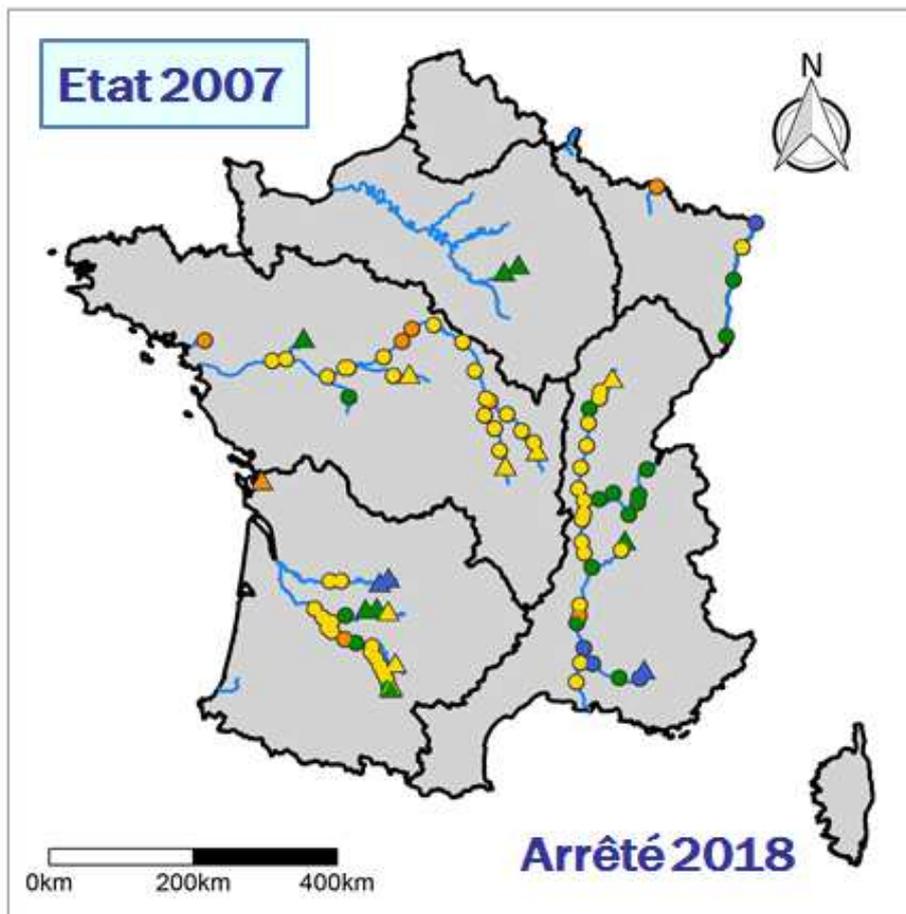
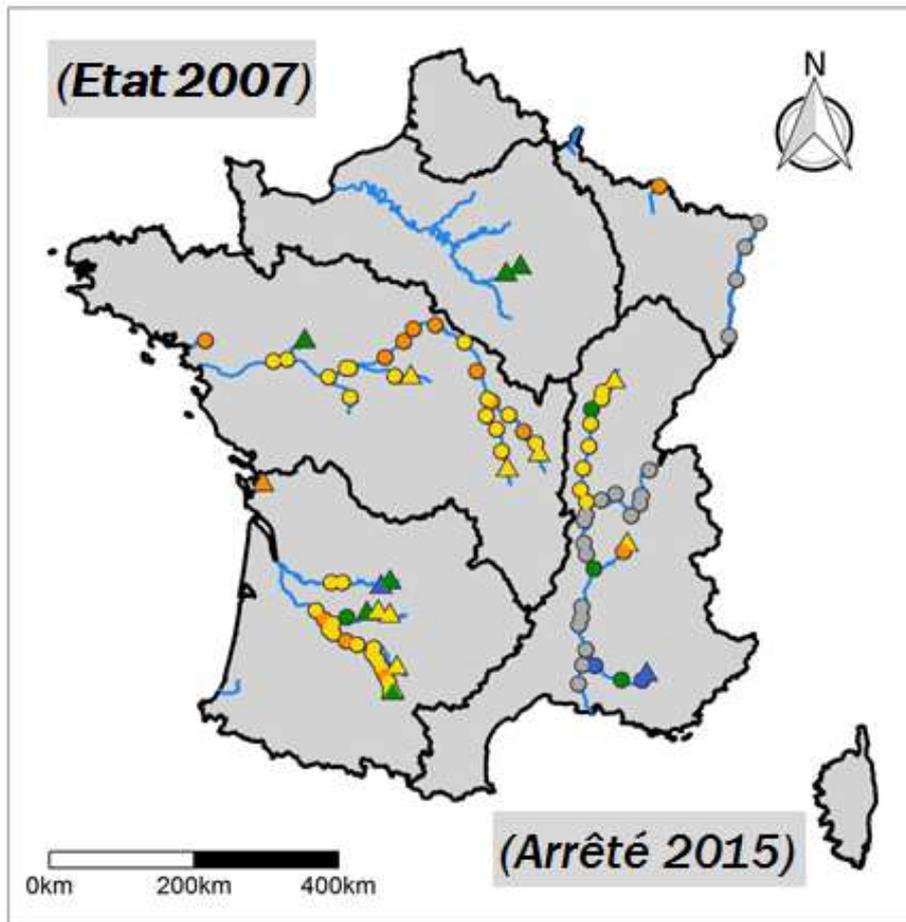
REF. 16 : Delmas, F. (2016). French Technical Note (08-05-2016). "Large Rivers GIG" : Remarks and questions about the report draft on preliminary results / BQE Phytobenthos (Version of 2016-05-03). Irstea Eds., 3 p.

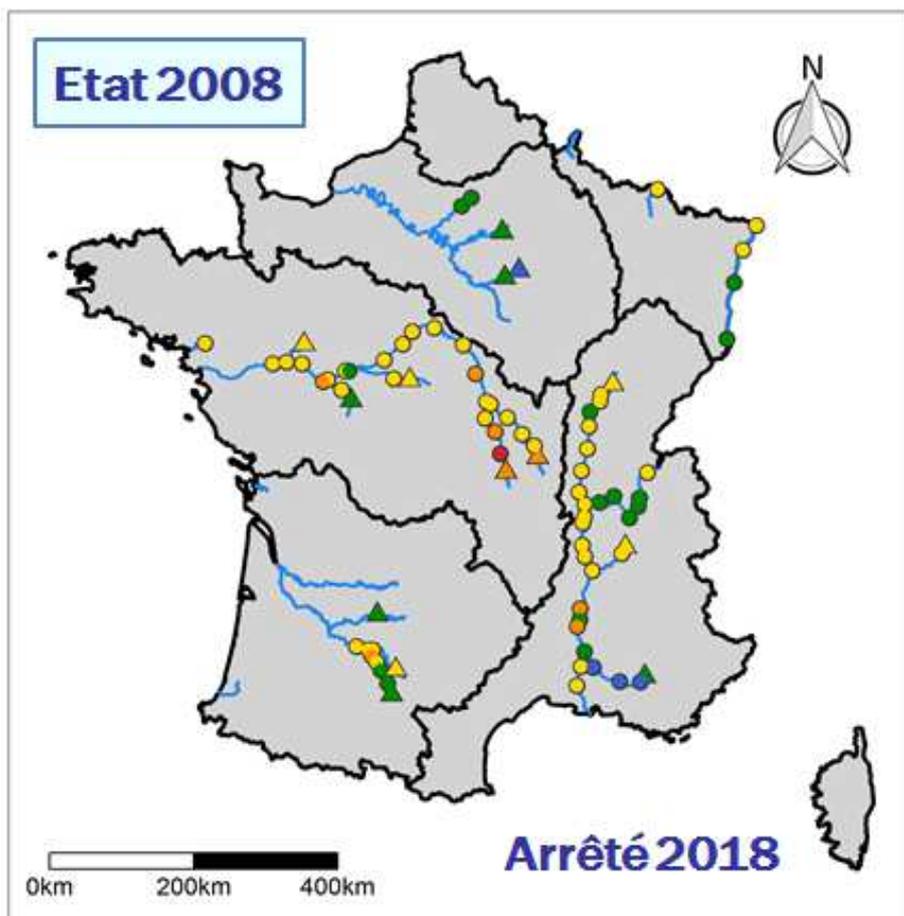
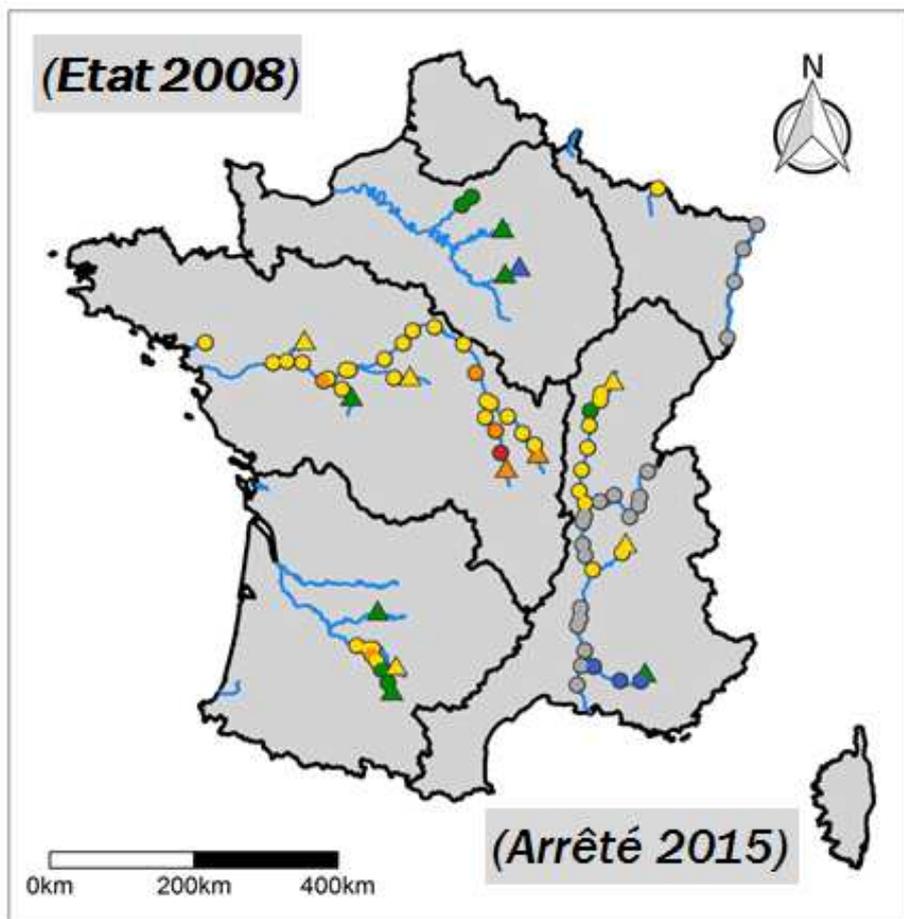
REF. 17 : Delmas, F. (2016). French Technical Note (16-05-2016). "Large Rivers GIG" : Complementary discussion elements about few methodological issues. Irstea Eds., 5 p.

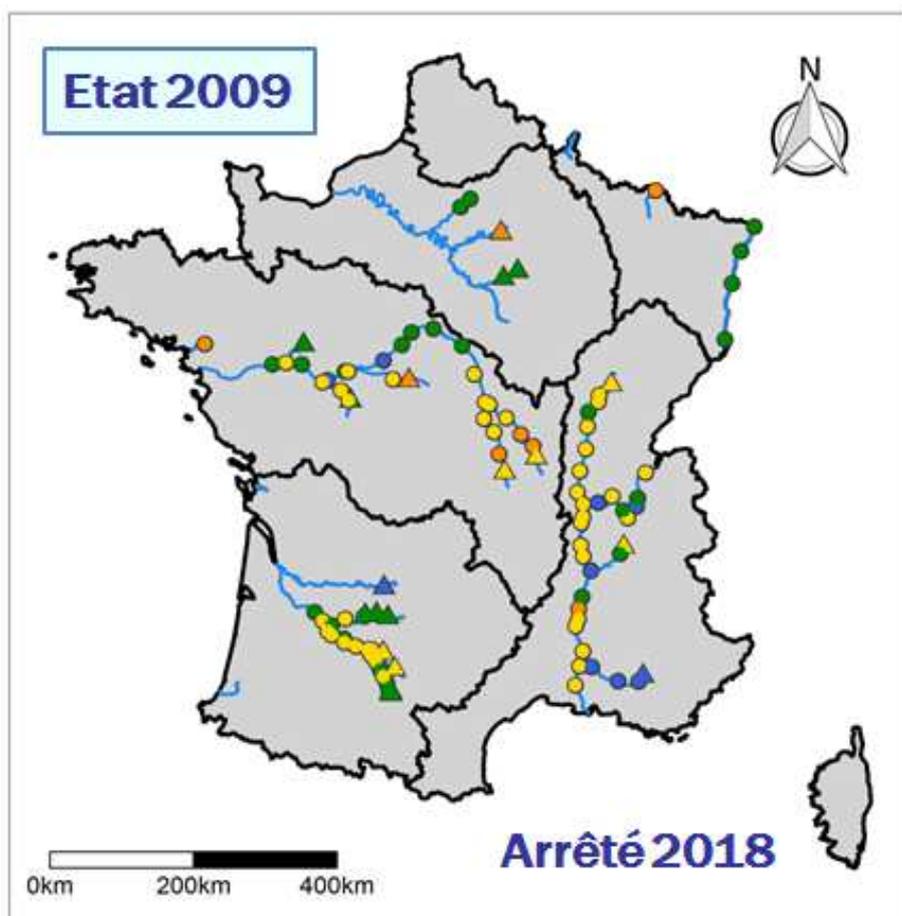
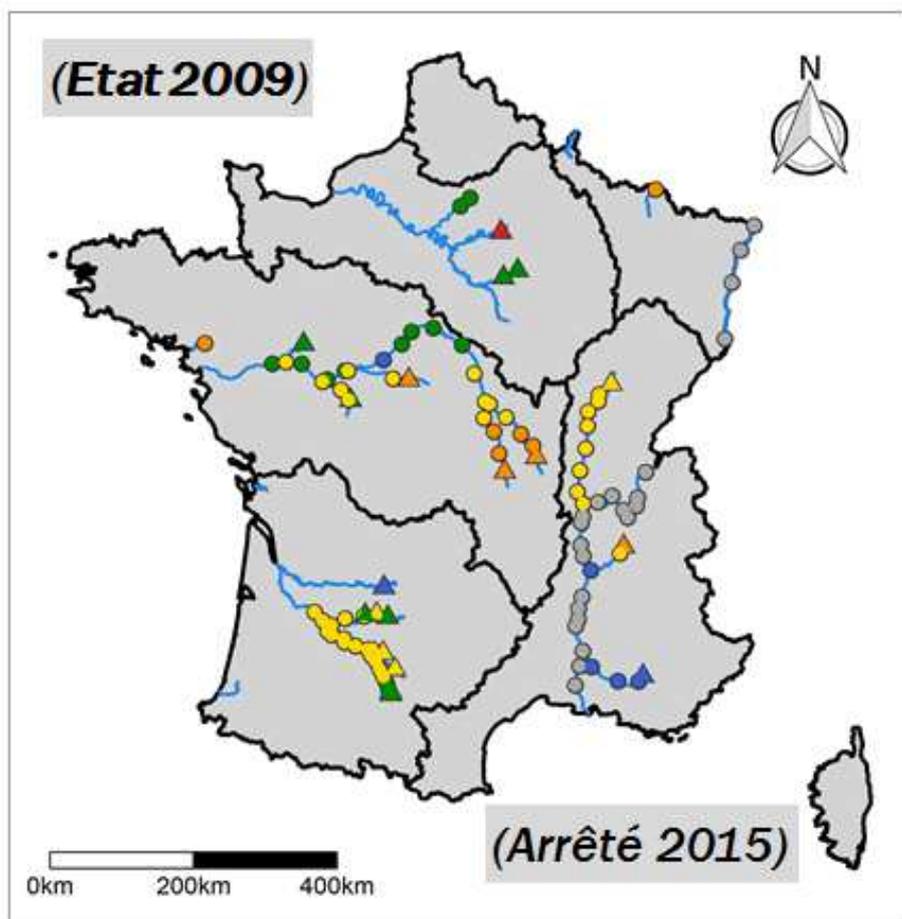
REF. 18 : Guéguen, J. & Delmas, F. (2016). French Technical and Expertise Note (03-06-2016). "Large Rivers GIG" : About intercalibration results of the 3rd report draft, version 2.1 (30-05-2016), and Scenario 2 results (02-06-2016). Irstea Eds., 10 p.

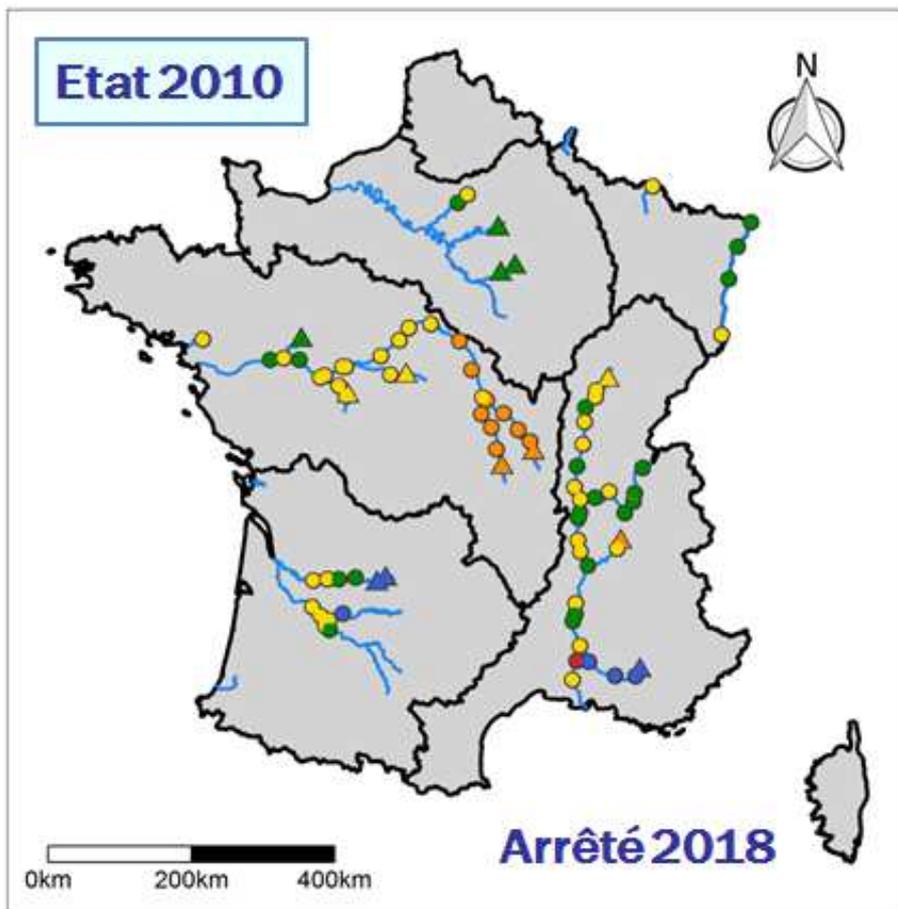
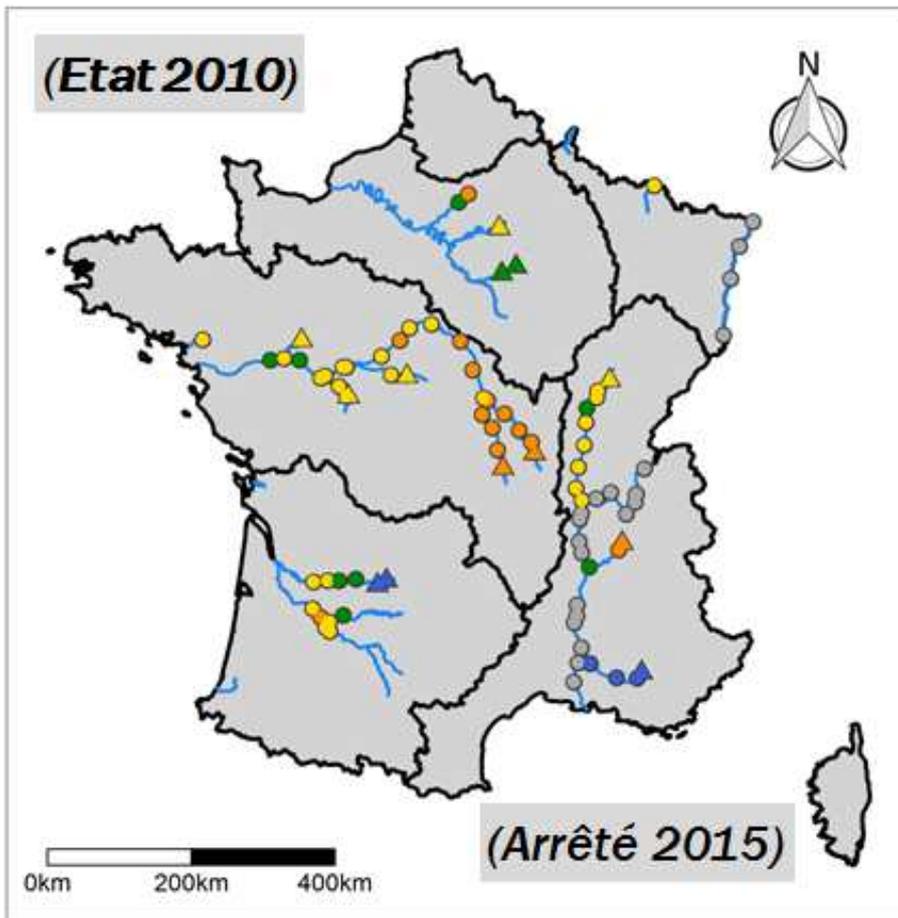
**ANNEXE 1 : Evaluation d'état diatomique des cours d'eau
candidats-TGCE de France (au seuil de 8 000 km² de BV intégré)**

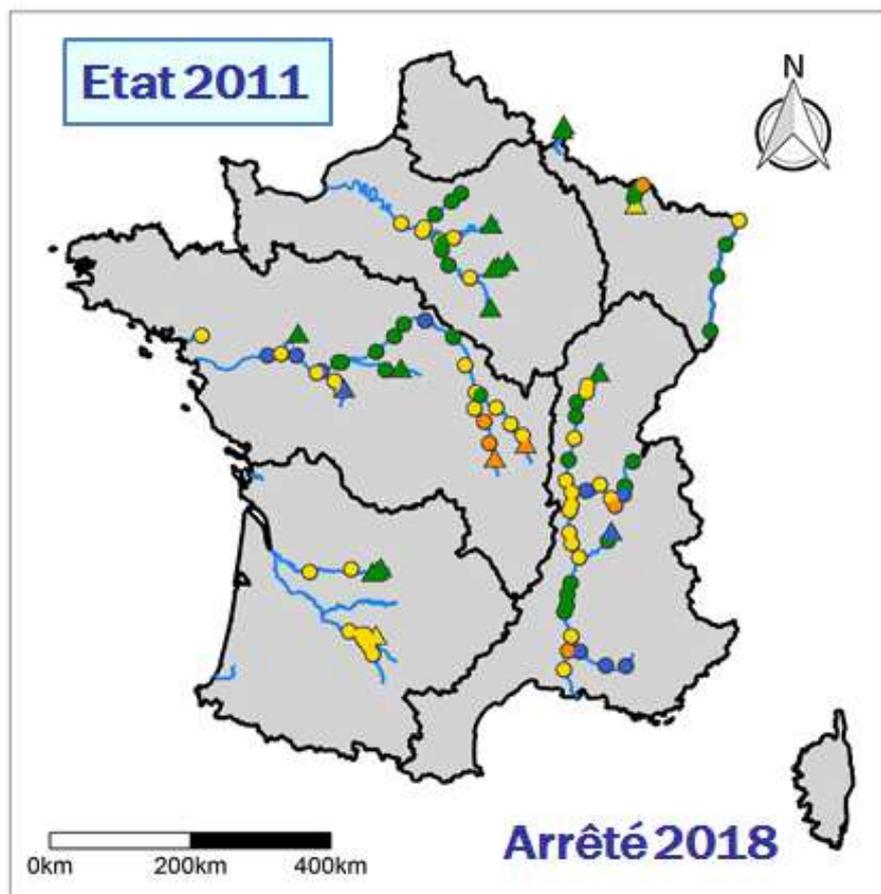
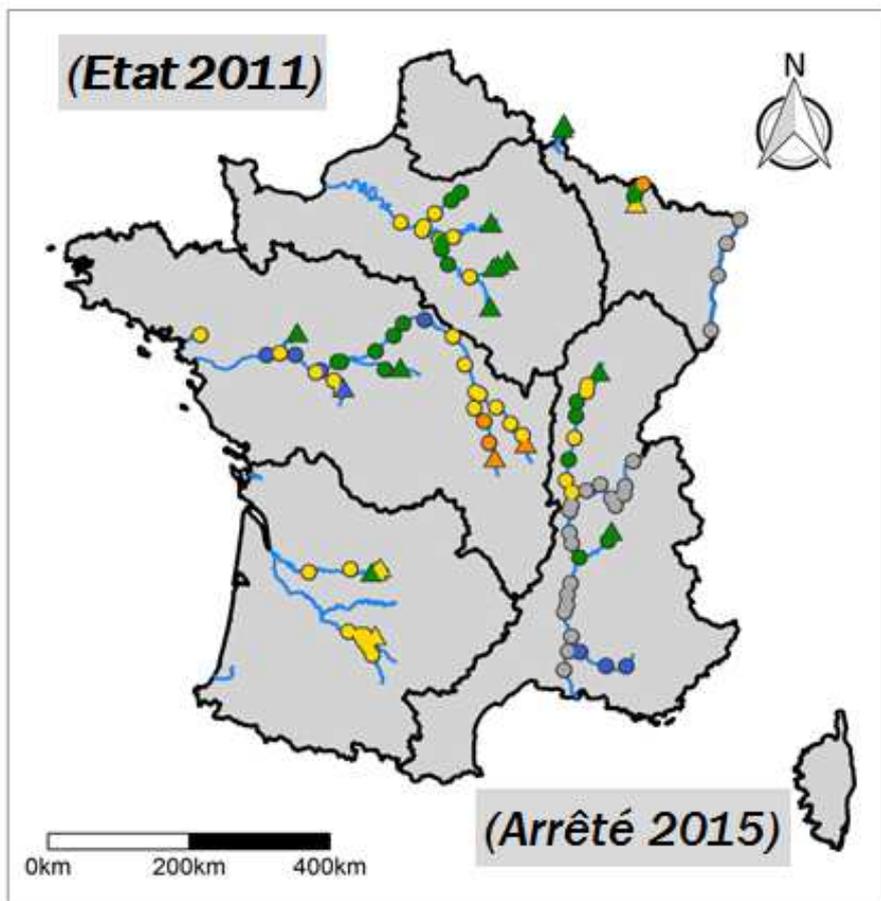
- En haut : selon dispositif antérieurement en vigueur (Arrêté «Evaluation» du 27 Juillet 2015)
- En bas : selon grilles spécifiques TGCE du nouvel Arrêté «Evaluation» du 27 Juillet 2018

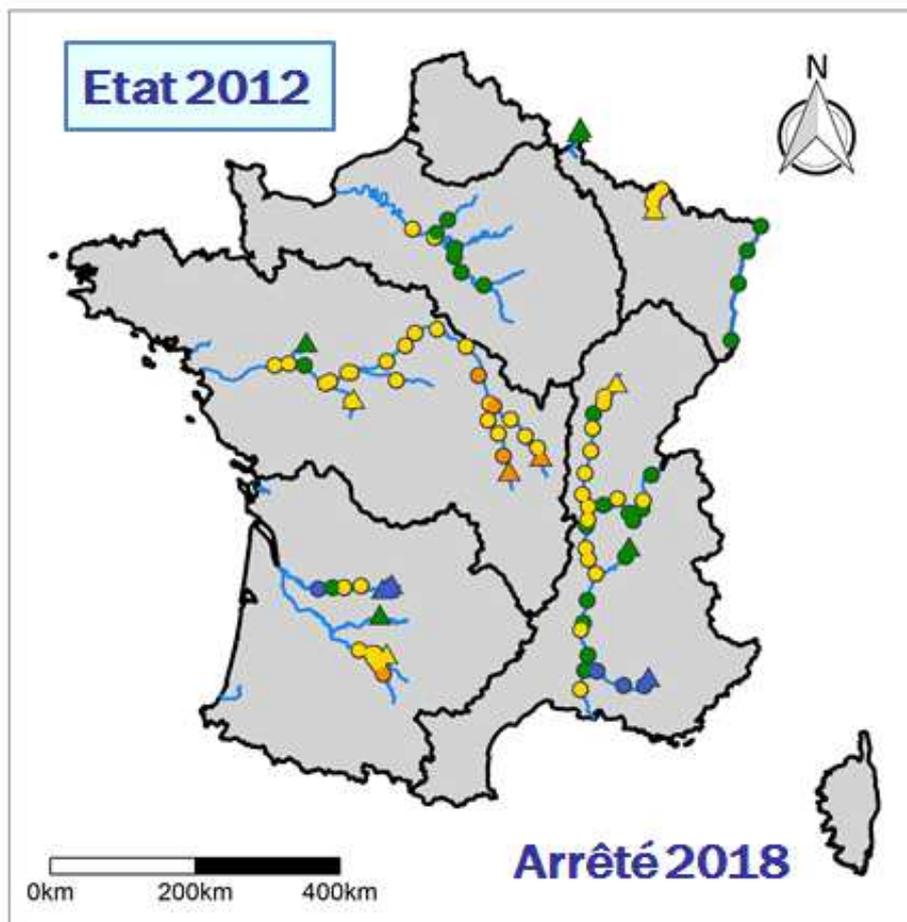
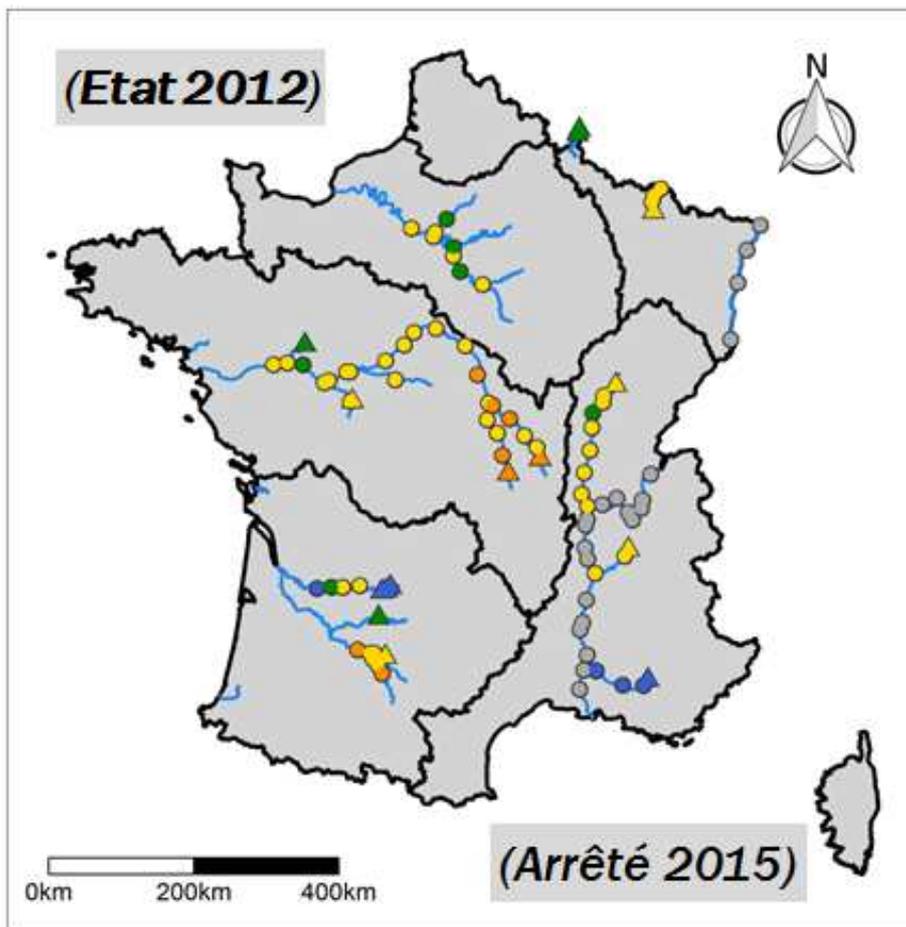


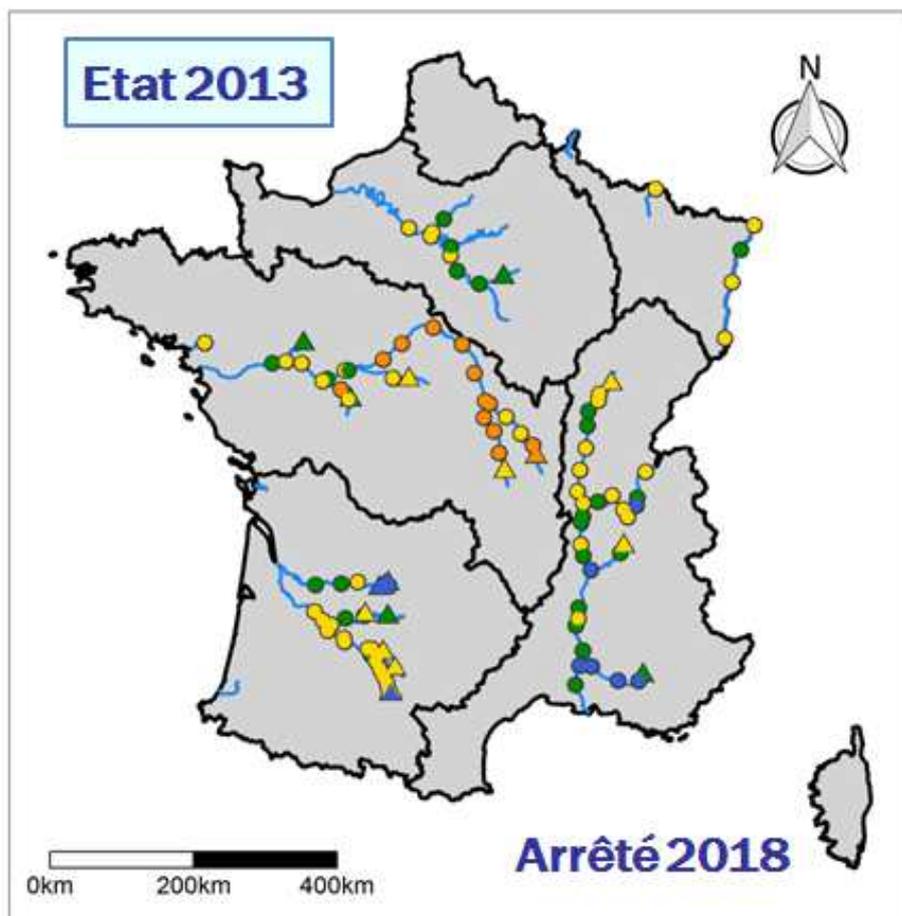
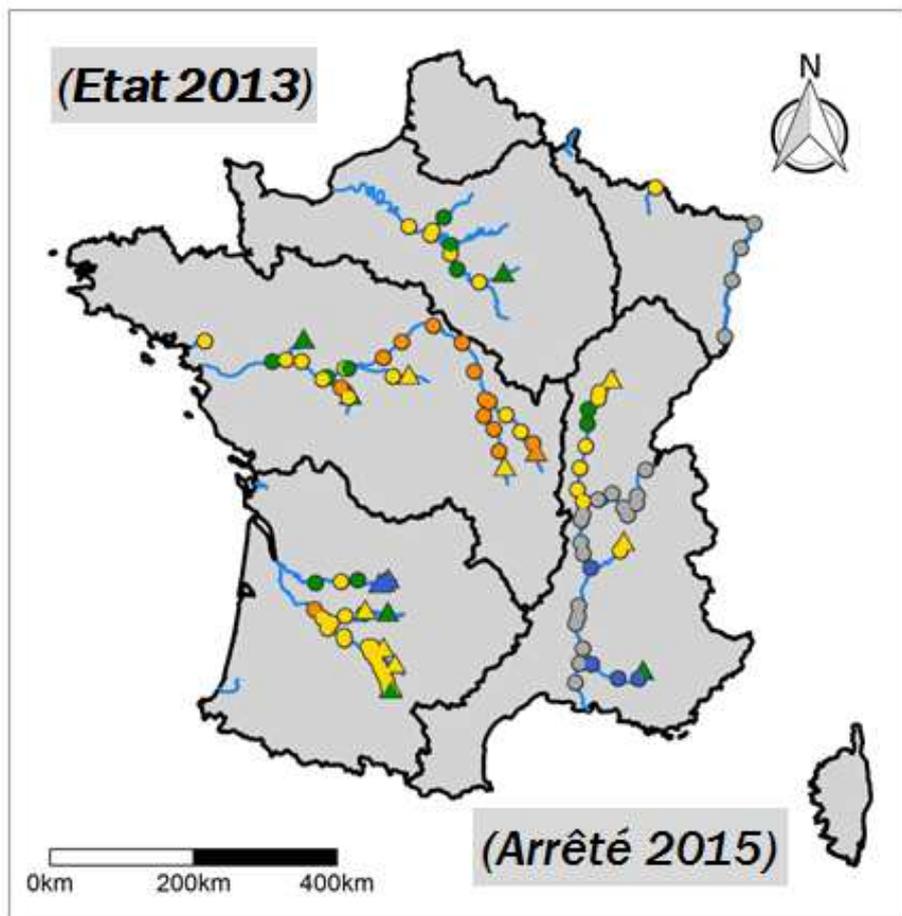












10. REMERCIEMENTS

Dans le cadre de la réalisation de cette étude nationale, il convient ici de remercier :

- les collègues d'IRSTEA Lyon / U.R. RIVERLY / Equipe LHQ impliqués de longue date dans la gestion et l'actualisation de la base de données Irstea « PANDORE » (notamment Martial Ferréol, Bertrand Villeneuve, André Chandesis...)
- les 6 Agences de l'Eau, ainsi que les collègues diatomistes des laboratoires Régionaux des DREAL, pour la fourniture des données abiotiques et diatomiques des réseaux de surveillance et de référence concernant leur bassin hydrographique,
- Nos collègues du Ministère en charge de l'Environnement / Direction de l'Eau (successivement Stéphane LUCET, Jean-Pierre CABARET et, pour les dernières phases de cette étude et les modificatifs d'Arrêté National, Benjamin JEANNOT), ainsi que Marina LE LOARER-GUEZBAR concernant plus spécifiquement le suivi et l'articulation de nos travaux dans le cadre des chantiers d'intercalibration européenne,
- et tout particulièrement l'AFB pour son soutien financier et pour l'implication de nos correspondants habituels, Yorick REYJOL et Bénédicte AUGÉARD (DAST), sans l'intervention desquels la réalisation de ce programme n'aurait pas été possible.

Irstea

UR EABX / Equipe CARMA,
Centre de Recherche de Bordeaux,
50, Avenue de Verdun,
33612 Cestas Cedex

05 57 89 08 00

www.irstea.fr

Agence Française pour la Biodiversité

Hall C – Le Nadar
5, square Félix Nadar
94300 Vincennes

01 45 14 36 00

www.afbiodiversite.fr