

Quelle limite de « bon état écologique » pour les invertébrés benthiques en rivières ? Apport des modèles d'extrapolation spatiale reliant l'indice biologique global normalisé à l'occupation du sol

Jean-Gabriel Wasson, Bertrand Villeneuve, Nicolas Mengin, Hervé Pella et André Chandèsris

Les macro-invertébrés benthiques sont les bioindicateurs les plus utilisés et sont à la base de l'indice biologique global normalisé (IBGN). Dans le cadre de la Directive cadre européenne sur l'eau, les auteurs présentent dans cet article des modèles d'extrapolation spatiale qu'ils ont développés au niveau du territoire français et qui permettent de relier l'état écologique des rivières, évalué à partir des bioindicateurs invertébrés, à l'occupation du sol. Deux hypothèses de bon état écologique sont développées et une première caractérisation qualitative des différentes zones sont discutées selon les pressions qui s'y exercent.

Contexte et objectifs

La Directive cadre européenne sur l'eau (DCE) place le vivant au cœur de la gestion des milieux aquatiques : l'objectif central assigné par cette directive est le « bon état écologique » qui se réfère « à la structure et au fonctionnement des écosystèmes aquatiques ». L'évaluation de cet état repose principalement sur des bioindicateurs – végétaux, invertébrés et poissons – qui seront en définitive les « juges de paix » (Wasson, 2002 ; Roche *et al.*, 2005). Le bon état est défini comme un écart « léger » par rapport à une situation de référence correspondant à des milieux très faiblement impactés par l'homme ; cet objectif s'impose pour toutes les rivières dont les caractéristiques physiques ne sont pas fortement modifiées pour un usage reconnu.

En réponse à une demande du ministère de l'Écologie et du Développement durable (MEDD), le Cemagref s'est fortement impliqué dans l'appui scientifique pour la mise en œuvre de la DCE : typologie des cours d'eaux, définition des valeurs de référence et révision des classifications pour les bioindicateurs les plus utilisés (invertébrés et diatomées). Cependant, les ajustements techniques sur les méthodes une fois réalisés, une question très concrète s'est posée aux décideurs : quelle valeur limite de « bon état écologique » peut, conformément aux exigences de la DCE, constituer un objectif à la fois « ambitieux et réaliste » selon les propres termes du directeur de l'Eau¹ ?

La réponse à cette question suppose, d'une part, de pouvoir simuler une application de différentes hypothèses à l'ensemble du territoire français pour en visualiser les conséquences, et d'autre part, de relier l'état écologique à des causes d'altération pour évaluer la faisabilité d'une amélioration.

Où se situent les problèmes ? Et sur quelles structures (urbaines, agricoles, industrielles...) faudrait-il agir en priorité ? Telles sont les questions corollaires à la fixation du niveau d'exigence de l'objectif de bon état écologique.

Ces questions nous ont orientés vers le développement de modèles « large échelle » pouvant fournir des réponses spatialisées, projetables sur des cartes. Elles soulèvent des problèmes scientifiques très actuels, à la fois en ce qui concerne les indicateurs et les modèles permettant une analyse spatiale, mais aussi les processus complexes qui déterminent les réponses biologiques aux pressions anthropiques (Allan, 2004), à différentes échelles d'espace et de temps.

En l'état des données disponibles sur les pressions anthropiques, il apparaît que seule l'occupation du sol permet de fournir des indicateurs pertinents et spatialement homogènes à l'échelle du territoire français. Comme indicateur biologique, nous avons choisi les invertébrés, avec l'indice IBGN² (encadré 1), robuste, largement utilisé et très comparable aux autres méthodes européennes.

1. Allocation prononcée au cours du séminaire « Bon état écologique », MEDD-DE, Paris, 26 octobre 2004.

2. Indice biologique global normalisé.

Les contacts

Cemagref, UR Biologie des écosystèmes aquatiques, Laboratoire d'Hydroécologie quantitative, 3 bis Quai Chauveau, CP 220, 69336 Lyon Cedex 09

Encadré 1

L'IBGN

L'IBGN, ou indice biologique global normalisé (AFNOR, 1992), est basé sur les peuplements de macro-invertébrés benthiques (Verneaux *et al.*, 1982). Il fait partie de la catégorie des indices biotiques développés en Europe de l'ouest depuis les années 60. Le protocole de prélèvement comprend huit relevés réalisés sur des substrats différents en cherchant à maximiser la diversité des habitats et la richesse faunistique. Les invertébrés sont identifiés au niveau de la famille. L'indice comprend deux métriques : la *variété taxonomique* (VT) ou nombre total de taxons exprimé en 14 classes, et le *groupe faunistique indicateur* (GFI) qui regroupe une quarantaine de taxons en 9 niveaux de sensibilité à la pollution. La note IBGN, donnée par la formule $(VT + GFI) - 1$, est une valeur entière comprise entre 1 et 20, la valeur 0 (absence totale de faune) étant exceptionnelle.

Les taxons indicateurs ont été sélectionnés plutôt en fonction de la pollution organique, mais ils s'avèrent aussi sensibles à différents types de contaminants chimiques et toxiques et à des dégradations de l'habitat (e.g. colmatage). De façon complémentaire, le nombre de taxons est sensible aux altérations de toute nature (physique, chimique, toxique) à partir du moment où elles sont suffisamment marquées pour éliminer plusieurs familles d'invertébrés. En revanche, les premiers stades de l'eutrophisation (par apport de nutriments) ou des modifications légères de l'habitat (e.g. éclaircissement) peuvent provoquer une augmentation de la richesse faunistique.

Bien que n'étant pas entièrement conforme aux préconisations de la DCE, l'IBGN est robuste et mesure une dégradation générale des peuplements en réponse à un large spectre de perturbations anthropiques. Il est actuellement comparé à d'autres méthodes européennes dans le cadre du projet STAR³ et de l'exercice d'inter-étalonnage requis par la DCE (Buffagni *et al.*, 2005).

Une analyse bibliographique préliminaire des travaux sur les relations à large échelle entre indicateurs biologiques et occupation du sol (Garcia

Bautista *et al.*, sous presse), ainsi que la revue de Allan (2004), ont montré qu'il existe relativement peu de publications sur le sujet. Il ne ressort pas de conclusion tranchée en ce qui concerne l'impact relatif de l'agriculture et de l'urbanisation sur l'état écologique des rivières, l'une ou l'autre cause apparaissant prépondérante selon les cas. Enfin, les seuils de pression exprimés en pourcentage d'occupation urbaine ou agricole du sol qui aboutissent à un risque élevé de dégradation des milieux sont très variables selon les contextes et les méthodes (Donohue *et al.*, 2006).

Nous avons donc développé des modèles « pressions/impacts » reliant l'état écologique des rivières évalué à partir des bioindicateurs invertébrés (IBGN) à l'occupation du sol. Ces travaux sont rassemblés dans un rapport qui expose le détail des méthodes et de nombreux résultats (Wasson *et al.*, 2005). Dans cet article, nous présentons les indicateurs utilisés, les principes d'un modèle d'extrapolation spatiale et son application à l'échelle du territoire français permettant de simuler deux hypothèses de « bon état écologique » pour l'IBGN. Cette recherche répond à la demande du MEDD et s'inscrit plus largement dans le projet européen REBECCA⁴.

Description des données

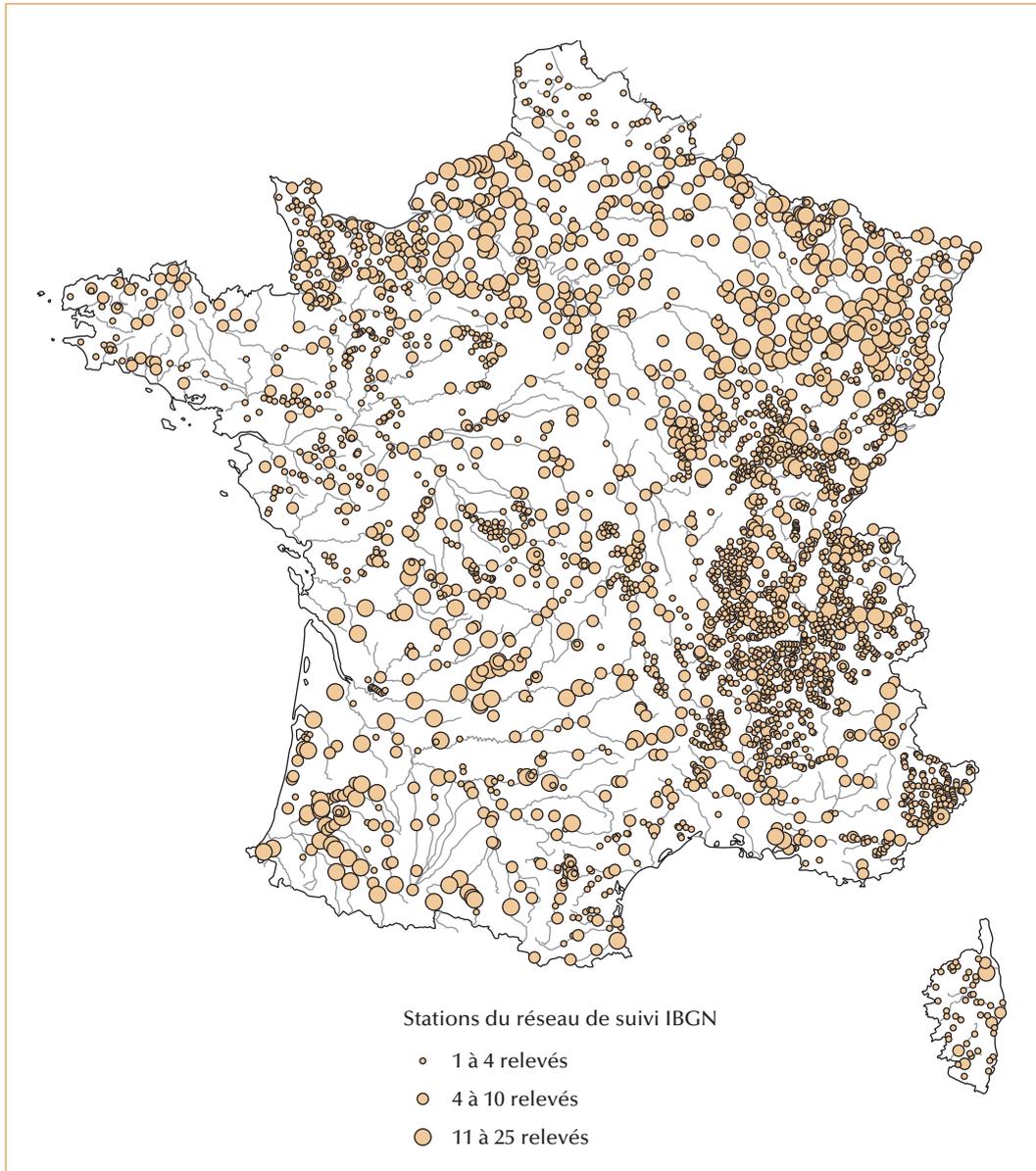
État écologique des stations d'après l'IBGN

Les invertébrés sont les bioindicateurs les plus utilisés actuellement au niveau européen, et leur place dans l'évaluation biologique pour la DCE sera probablement centrale. L'état écologique des stations est ici évalué à partir de l'IBGN utilisé depuis 1992 dans les réseaux de suivi. Néanmoins, il faut souligner que l'IBGN ne saurait définir à lui seul l'état écologique d'une station, qui devra nécessairement inclure l'évaluation d'autres peuplements (diatomées, poissons, macrophytes) soutenus par les paramètres physico-chimiques. Pour éviter toute confusion, nous emploierons le terme « état écologique IBGN » dans la suite de ce travail.

Nous avons constitué une base de données nationale des stations biologiques suivies par les directions régionales de l'Environnement (DIREN). Cette base comporte actuellement 3 662 stations pour lesquelles on dispose de 12 682 relevés (ou notes IBGN) couvrant la période 1992-2002 (figure 1). Près de 70 % des

3. Cf. <http://www.eu-star.at>

4. Cf. <http://www.environment.fi/syke/rebecca>



◀ Figure 1 – Représentation cartographique des stations du réseau national de suivi IBGN.

relevés sont accompagnés d'une liste faunistique des peuplements.

L'état écologique IBGN de chacune des stations a été déterminé selon la procédure suivante en quatre étapes :

- **Typologie** – Chaque station est affectée à un type de cours d'eau en fonction de sa position géographique. La typologie est basée principalement sur les hydro-écorégions (Wasson *et al.*, 2002, 2004a) et sur une classe de taille déter-

minée par le rang de Strahler du cours d'eau (Chandesris *et al.*, 2006).

- **Référence** – À chaque type de cours d'eau est affectée une note IBGN de référence. Cette note de référence correspond à la médiane des valeurs IBGN observées sur des sites de référence. Ces sites ont été sélectionnés selon une double procédure (Wasson *et al.*, 2004b), cohérente avec les préconisations du guide européen REFCOND (Wallin *et al.*, 2003), comportant d'une part un avis d'expert de terrain (DIREN), d'autre part

5. Système d'évaluation de la qualité des cours d'eau (Agences de l'eau, 1999).

6. Valeur pour laquelle 25 % des échantillons mesurés sont inférieurs.

▼ Figure 2 – Hypothèses de limite de bon état : limite « basse » d'après une suggestion du guide REFCOND, et limite « haute » relevée d'un point IBGN (M = mauvais état écologique).

une évaluation sous SIG (système d'information géographique) des paramètres d'occupation du sol et autres pressions.

• **Écart à la référence** – Les notes IBGN sont alors normées par la valeur de référence de chaque type et transformées en « écart à la référence » ou EQR (pour *ecological quality ratio*), variant entre 0 pour l'altération maximale et des valeurs autour de 1 en situation de très bon état. Cet EQR-IBGN, utilisé dans toutes les analyses suivantes, permet donc de déterminer un état écologique IBGN indépendamment des variations naturelles (typologiques) de l'indice.

• **Classement d'une station** – Il faut enfin positionner chaque station dans une classe d'état écologique. Pour tenir compte de la variabilité naturelle de la faune et de l'erreur inhérente à toute méthode, le classement est basé sur la distribution des IBGN observés sur chaque station. Suivant la logique utilisée dans le SEQ Eau⁵ pour évaluer la qualité physico-chimique à partir du percentile 90 d'une chronique de données (Oudin et Maupas, 2003), nous avons adopté la règle suivante : « Pour positionner une station dans une classe donnée, il faut que les 3/4 au moins des valeurs IBGN observées soient supérieures ou égales à la limite inférieure de la classe ». Cette règle de classification est sensiblement plus exigeante que l'utilisation de la moyenne.

Hypothèses de limite du « bon état » sur l'IBGN

L'état écologique se décline en 5 classes, mais les objectifs environnementaux de la DCE (article 4) concernent uniquement le « très bon état » (TBE) et le « bon état » (BE), dont les limites sont actuel-

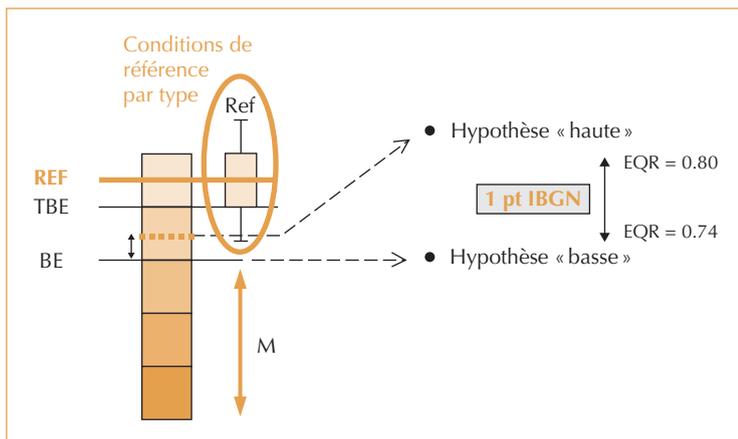
lement harmonisées entre les États membres au cours de l'exercice d'inter-étalonnage ; les autres classes sont définies au niveau national. Le très bon état correspond plus ou moins à la gamme de variabilité naturelle des peuplements. La limite TBE a donc été fixée pour chaque type, suivant une recommandation du guide communautaire européen REFCOND, au 25^e percentile⁶ de la distribution des valeurs IBGN observées dans les sites de référence.

Pour les invertébrés des rivières, selon l'annexe V de la DCE, la principale définition normative du bon état écologique concerne un écart « léger » aux conditions de référence sur différentes métriques (composition taxonomique, abondance et diversité, taxons sensibles). À l'évidence, une telle définition laisse une certaine marge d'interprétation. En première hypothèse, et suivant une méthode proposée par le guide REFCOND, nous avons défini une limite de bon état qui revient à diviser en 4 parties égales l'intervalle entre la limite TBE et la valeur minimum de l'IBGN. En moyenne pour l'ensemble des types, cette limite correspond à un EQR-IBGN de 0,74.

Cependant, les premiers tests d'application de cette hypothèse (Wasson *et al.*, 2005) ont montré que cette limite ne discriminait que des situations déjà assez dégradées. Cette constatation rejoignait le sentiment exprimé par certains acteurs de terrain qu'elle ne correspondait pas à leur perception du « bon état ». Nous avons donc étudié une deuxième hypothèse en remontant uniformément d'un point IBGN (soit 6 % d'EQR-IBGN) la limite BE calculée en première approche, à l'exception de quelques types de cours d'eau (Alpes, Cévennes...) pour lesquels, en raison de caractéristiques naturelles, cette option ne semblait pas réaliste. Cette hypothèse « haute » correspond à un EQR-IBGN de 0,80 en moyenne pour l'ensemble des types. Nous étudions dans ce travail ces deux hypothèses, que nous appellerons « haute » et « basse », de limite de BE pour l'IBGN (figure 2).

Les indicateurs de pressions anthropiques

Pour évaluer le degré d'anthropisation des cours d'eau de façon cohérente et homogène à l'échelle du territoire national, la seule possibilité au vu des données existantes est d'analyser l'occupation du sol, représentative des « structures anthropiques » exerçant des pressions sur les milieux aquatiques.



L'occupation du sol a donc été calculée pour le bassin versant de chaque station IBGN. La couche d'information géographique utilisée est Corine Land Cover⁷ (CLC) 2000, établie à partir d'images satellitales Landsat et Spot. Elle couvre l'ensemble du territoire et repose sur une nomenclature standard hiérarchisée à 3 niveaux et 44 catégories réparties selon 5 grands types d'occupation du territoire (territoires artificialisés, territoires agricoles, forêts et milieux semi-naturels, zones humides et surfaces en eau).

Le couplage entre les données biologiques et les pressions s'exerçant sur le bassin des stations a nécessité des développements informatiques spécifiques sous SIG (Pella *et al.*, 2004) pour exploiter les couches d'informations issues de l'IGN (Institut géographique national) : modèle numérique de terrain (MNT) dérivé de la BD ALTI[®] au pas de 250 m, découpage en zones hydrographiques (ZH) et tracé du réseau hydrographique tirés de la BD Carthage[®] (version 3.0). Une fois la station raccrochée géographiquement sur le tracé du cours d'eau auquel elle appartient, le traitement consiste à découper le MNT en fonction de l'enveloppe du bassin amont définie à partir des zones hydrographiques, puis à déterminer son bassin versant réel. Le pourcentage de surface des différentes catégories d'occupation du sol Corine Land Cover a été calculé à partir du bassin versant ainsi délimité.

Modèles d'extrapolation spatiale

Pour établir des modèles pressions/impacts, plusieurs étapes techniques sont nécessaires : analyse des données de pressions, sélection des variables pertinentes vis-à-vis de la cible biologique, mise en œuvre de méthodes prédictives permettant d'extrapoler les résultats des modèles sur une carte (Wasson *et al.*, 2005). Dans ce travail, nous avons utilisé des arbres de décision pour construire un modèle d'extrapolation spatiale. L'algorithme complet de construction des arbres de décision est décrit dans Breiman *et al.* (1984). En sortie de modèle, l'arbre de décision prédit pour la variable cible une classe (variable qualitative) ou une valeur numérique (variable quantitative). Dans le cas de l'état écologique IBGN, la variable cible est qualitative : on cherche à prédire l'atteinte de l'objectif DCE, que l'on appellera pour simplifier « bon état » pour les stations classées BE ou TBE, et « mauvais état » pour les autres (IBGN inférieur à la limite BE).

Dans ce cas, l'évaluation de l'efficacité du modèle se fait à partir de la matrice de confusion. Dans un premier temps, cette matrice permet de calculer le taux de mauvaise classification du modèle, qui constitue une première estimation de l'erreur (ou taux d'erreur du modèle). Une courbe ROC (*receiver operating characteristic*, Hanley et McNeil, 1982) est ensuite utilisée comme « courbe de caractéristique d'efficacité ». L'aire sous cette courbe ROC (AUC : *area under curve*) est un estimateur de l'efficacité globale du test. Lorsque l'aire est de 0,5, le test n'est pas informatif, et si l'aire est de 1, le test est parfaitement discriminant.

Sélection des classes simplifiées d'occupation du sol

Pour développer un modèle prédictif en arbre de décision, il est nécessaire de réduire le nombre de variables d'entrée, car un trop grand nombre de variables peut entraîner une instabilité du modèle et des difficultés à le valider. Cette sélection a été réalisée par une analyse de régression PLS (*partial least square*) entre l'EQR-IBGN et l'occupation du sol des bassins (Wasson *et al.*, 2005). Les quatre classes retenues sont les suivantes :

- **territoires artificialisés** : zones urbanisées, zones industrielles ou commerciales et réseaux de communication, mines, décharges et chantiers, espaces verts artificialisés, non agricoles – codes CLC : 1.1.1, 1.1.2, 1.2, 1.3, 1.4.1, 1.4.2 ;
- **territoires agricoles à fort impact** : terres arables, cultures permanentes, vergers, vignes, oliveraies, cultures annuelles associées aux cultures permanentes, systèmes culturaux et parcellaires complexes – codes CLC : 2.1, 2.2, 2.4.1, 2.4.2 ;
- **territoires agricoles à faible impact** : prairies, territoires principalement occupés par l'agriculture avec présence de végétation naturelle importante, territoires agro-forestiers – codes CLC : 2.3.1, 2.4.3 ;
- **territoires à faible anthropisation** : forêts et milieux semi-naturels, zones humides, surfaces en eau – codes CLC : 3.1.1, 3.1.2, 3.1.3, 3.2, 3.3, 4 et 5.

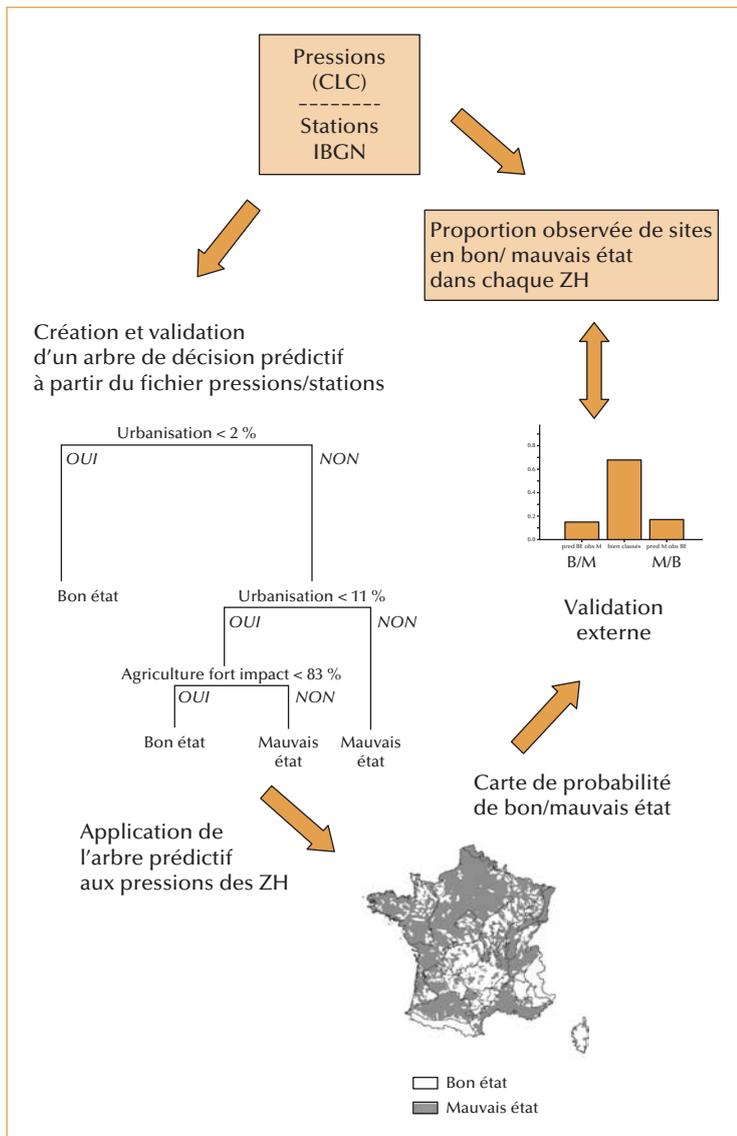
Construction et validation des modèles

Le développement des modèles d'extrapolation spatiale se déroule selon le protocole décrit dans la figure 3. Les 3 662 stations IBGN sont regroupées en deux catégories selon leur état écologique IBGN « bon » ou « mauvais » comme défini précédemment.

7. Cf. <http://image2000.jrc.it/index.html>. Les images Spot correspondantes datent de 1998-99, soit au milieu de la période des relevés IBGN (1992-2002).

En premier lieu, un arbre de décision est développé et validé à partir de l'état écologique IBGN observé des 3 662 stations, et de l'occupation du sol, exprimée en quatre classes, des bassins versants réels de ces stations. Cet arbre de décision est ensuite utilisé pour prédire l'état écologique IBGN probable (prédit « bon » ou « mauvais ») des drains principaux des zones hydrographiques (ZH), d'après l'occupation du sol du bassin versant cumulé à l'amont de chaque ZH.

▼ Figure 3 – Description du processus de modélisation et d'extrapolation spatiale de l'arbre de décision validé. Le modèle et la carte présentés correspondent aux résultats du modèle « hypothèse haute » de bon état.



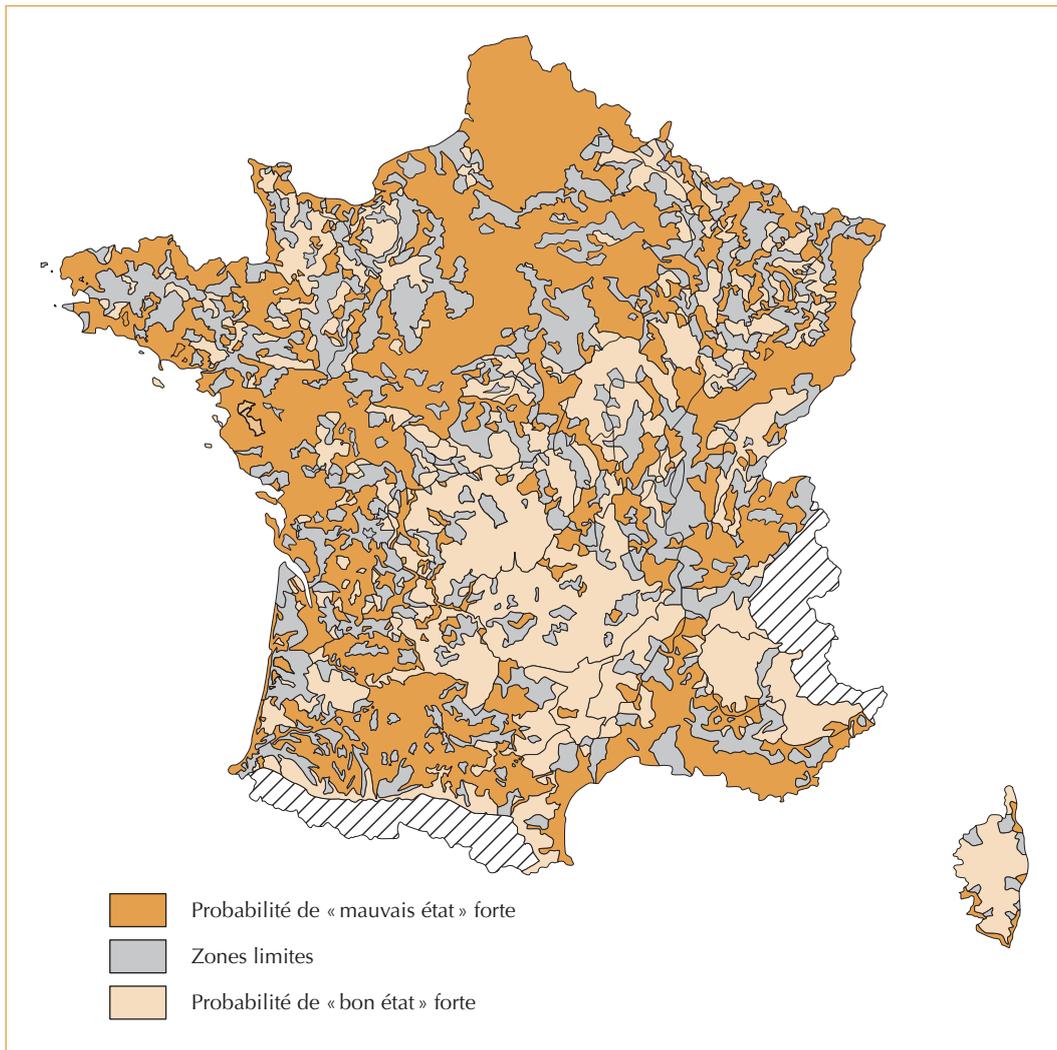
On utilise les sorties prédictives du modèle pour représenter sur une carte l'état écologique IBGN probable (prédit) pour le drain principal de chaque ZH. On procède enfin à une validation externe de la représentation cartographique en comparant l'état des ZH « prédit » par le modèle et l'état « observé » sur des stations situées sur les drains principaux des mêmes ZH. L'intérêt de ce type de modèle est de pouvoir tester différentes hypothèses de classifications et d'en visualiser le résultat à l'échelle de la France entière. Nous avons testé la réponse des modèles dans les deux hypothèses « haute » et « basse » de limite du bon état IBGN définies précédemment. L'objectif est de visualiser les conséquences de chaque hypothèse en termes de probabilité d'atteinte du bon état dans les conditions actuelles, et d'identifier les zones en situation limite.

Résultats

Modèles d'extrapolation spatiale de l'état écologique IBGN

Le modèle prédictif pour la France entière de l'état écologique IBGN des stations à partir des quatre classes d'occupation du sol, avec l'hypothèse « haute » de limite BE, donne un arbre de décision validé de 23 nœuds. Ce modèle est d'une efficacité prédictive satisfaisante ($AUC = 0,75$) et son erreur de prédiction est de 29 % (3 chances sur 10 de se tromper en prédisant l'état écologique du drain principal d'une ZH). Le même modèle avec l'hypothèse « basse » de limite BE donne un arbre de décision validé de 45 nœuds, avec une efficacité également satisfaisante ($AUC = 0,77$) et une erreur de prédiction de 27 %.

Ces modèles ont été utilisés pour prédire l'état IBGN actuel probable « bon » ou « mauvais » des zones hydrographiques, d'après l'occupation du sol du bassin versant cumulé à l'amont de chaque ZH. On utilise les sorties prédictives pour représenter sur une carte la probabilité actuelle de non atteinte du bon état IBGN pour le drain principal de chaque ZH dans les deux hypothèses (figure 4). Cette carte correspond à la représentation simultanée des deux modèles : les ZH prédites en « mauvais état » dans les deux cas sont figurées en brun foncé ; celles qui sont prédites en « bon état » dans les deux cas sont en brun clair ; enfin les ZH prédites en « bon état » dans l'hypothèse basse et en « mauvais état » dans l'hypothèse haute sont figurées en gris. Ces



◀ Figure 4 – Représentation cartographique de la probabilité actuelle de bon état écologique IBGN d'après le modèle d'extrapolation spatiale. Le modèle ne s'applique pas dans les zones de haute montagne (Alpes et Pyrénées).

zones en « situation limite » se situent essentiellement dans les hydro-écorégions de plaine à forte activité agricole : Tables Calcaires, Massif Armoricaïn, Coteaux Aquitains, Plaine de Saône, Méditerranée... Des tests préliminaires (Wasson et al., 2005) ont montré que ce type de modèle n'était pas applicable dans les hydro-écorégions de haute montagne, pour lesquelles l'erreur est fortement biaisée ; l'extrapolation n'a donc pas été réalisée dans ces zones (Alpes internes et Pyrénées) figurées en trame hachurée.

La validation externe (tableau 1) montre que le modèle de l'hypothèse basse est plutôt « optimiste » (erreur déplacée dans le sens d'une prédiction de bon état), alors que le modèle avec l'hypothèse haute est plutôt « pessimiste »

(erreur déplacée dans le sens d'une prédiction de mauvais état). Les modèles prédictifs amplifient donc légèrement la tendance observée dans les

▼ Tableau 1 – Validation externe des modèles. Comparaison de l'état des sites observés dans le réseau de mesure et de leur état prédit par le modèle d'extrapolation en fonction des deux hypothèses testées de limite de bon état.

	État des zones hydrographiques (ZH)		
	BE prédit ME	Bien prédit	ME prédit BE
Hypothèse haute	21 %	66 %	13 %
Hypothèse basse	12 %	67 %	21 %

données du réseau ; ce biais porte sur 8 à 9 % des ZH, ce qui revient à exagérer la tendance d'environ 4 % dans un sens ou dans l'autre. Il convient donc d'en tenir compte dans l'interprétation de la figure 4.

Comparaison des deux hypothèses de limite de bon état IBGN

L'image fournie par le réseau peut également être biaisée par le fait que les stations sont généralement placées à des points de contrôle de la pollution. L'image extrapolée par le modèle corrige ce biais en appliquant les mêmes critères à l'ensemble des drains principaux des ZH. Il est donc instructif de comparer la proportion de stations classées en « bon » ou « mauvais » dans les réseaux de suivi, avec les résultats des modèles (tableau 2).

L'image extrapolée sur les ZH a tendance à amplifier les proportions BE/ME observées sur

les stations du réseau de suivi. Ainsi, avec la limite BE haute, il y aurait 38 % seulement de ZH en bon état selon le modèle, contre 49 % des stations dans le réseau de suivi. À l'inverse, avec l'hypothèse basse, il y aurait 64 % de ZH en bon état avec le modèle, contre 60 % dans le réseau de suivi. Les modèles semblent donc corriger le biais lié à la situation des points du réseau de suivi.

Cependant, il faut également tenir compte du décalage de l'erreur lié aux modèles eux-mêmes (tableau 1). Autrement dit, la proportion des stations en bon état fournie par les résultats bruts des modèles est sous-estimée dans l'hypothèse haute, et surestimée dans l'hypothèse basse, le décalage étant d'environ 4 % dans les deux cas. Si l'on « corrige » les résultats bruts des modèles en retranchant ce décalage, on peut estimer la proportion des stations en « bon état IBGN » dans ces différentes hypothèses (tableau 3).

Les résultats sont concordants pour l'hypothèse basse, avec 60 % de stations en bon état dans le réseau et selon les prédictions corrigées des modèles. Mais dans l'hypothèse haute, il y aurait 42 % de ZH en « bon état » selon le modèle corrigé, contre 49 % dans le réseau. Après correction du modèle, le point important est la proportion élevée (18 %) de ZH en « situation limite », à un point IBGN seulement en dessous de la limite du bon état.

Influence de l'occupation du sol sur l'état écologique IBGN

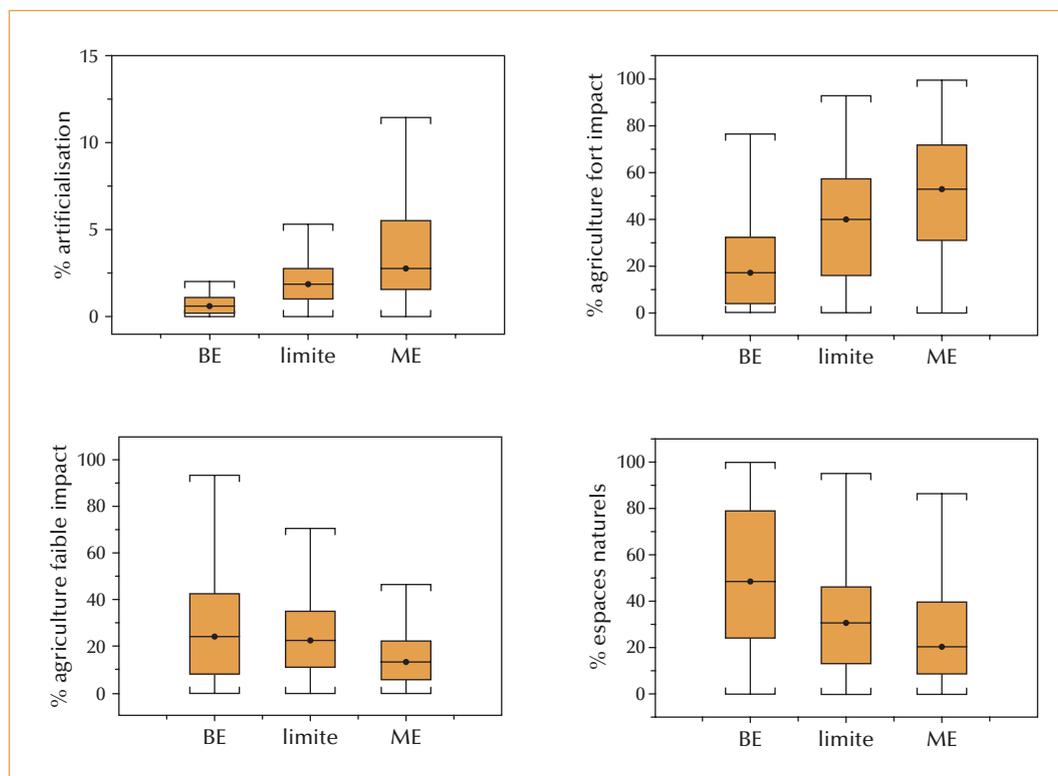
La figure 5 permet de visualiser l'occupation du sol dans les bassins amont des zones hydrographiques classées par les modèles en « bon état » ou en « mauvais état » dans les deux hypothèses, et les zones en « situation limite » classées en bon état dans l'hypothèse basse et en mauvais état dans l'hypothèse haute. La tendance croissante du pourcentage d'espaces artificialisés et de l'agriculture à fort impact (terres labourées) du « bon état » vers le « mauvais état » indique un effet clairement négatif de ces catégories d'occupation du sol sur l'IBGN. Cependant, on notera que l'impact des territoires artificialisés intervient pour des pourcentages d'occupation du sol beaucoup plus faibles (de 0 à 12 %) en comparaison de l'agriculture (de 0 à 100 %), ce qui indique un fort impact des zones urbanisées. L'agriculture à faible impact (prairies) semble avoir un effet

▼ Tableau 2 – État écologique IBGN des zones hydrographiques prédit par le modèle d'extrapolation spatiale et observé dans les stations du réseau de suivi, selon les deux hypothèses de bon état simulées.

	État des ZH prédit par les modèles		État observé des stations du réseau de suivi	
	BE	ME	BE	ME
Hypothèse haute	38 %	62 %	49 %	51 %
Hypothèse basse	64 %	36 %	60 %	40 %

▼ Tableau 3 – État écologique IBGN actuel probable des zones hydrographiques et des stations du réseau de suivi. Proportion de ZH ou de stations en « bon état » et en « situation limite » selon les deux hypothèses simulées, en fonction des prédictions des modèles, des prédictions corrigées et des données des réseaux.

	ZH ou stations en « Bon état IBGN »		ZH ou stations en « situation limite »
	Hypothèse haute	Hypothèse basse	
Modèles bruts	38 %	64 %	26 %
Modèles « corrigés »	42 %	60 %	18 %
Réseau de suivi	49 %	60 %	11 %



▲ Figure 5 – Boîtes à moustaches représentant les distributions des variables d’occupation du sol amont des zones hydrographiques en fonction de leur classe en sortie du modèle d’extrapolation spatiale. Le seuil limite d’artificialisation pour le bon état est estimé à 2 % et le seuil limite d’agriculture à fort impact pour le bon état est estimé à 80 %.

relativement limité sur l’appartenance des ZH à une classe donnée en sortie des modèles. Par contre les espaces naturels sont beaucoup plus représentés dans les bassins des ZH classées en « bon état » que dans les autres, ce qui illustre un effet clairement positif.

Ces résultats permettent aussi d’identifier des seuils d’occupation du sol compatibles avec un « bon état IBGN ». Les modèles ne prédisent aucune ZH en bon état (hypothèse haute) au-dessus d’un seuil de 2 % de territoires artificialisés dans les bassins versants, et les 3/4 des ZH prédites en bon état ont moins de 1,25 % d’urbanisation. Pour l’agriculture à fort impact (terres arables et cultures permanentes), le seuil maximum pour une prédiction de bon état est de 80 %, ce qui est assez élevé, mais on note aussi que les 3/4 des ZH prédites en BE ont moins de 30 % de cette catégorie d’occupation du sol dans leur bassin.

Discussion

Intérêt et limites des modèles d’extrapolation spatiale

Le principal intérêt de ces modèles réside dans la possibilité de tester différentes hypothèses de classification à l’échelle nationale, en s’affranchissant des biais qui peuvent être liés à la répartition des points dans les réseaux de mesure. Ils permettent également de visualiser les résultats pour donner une image globale de l’état actuel des milieux. Toutefois, il faut garder en mémoire que l’extrapolation est limitée aux drains principaux des zones hydrographiques et ne représente pas l’état des petits affluents. L’erreur locale des modèles est relativement importante, mais faiblement déséquilibrée ; il ne faut donc pas chercher à interpréter l’état d’une ZH particulière (le risque d’erreur est alors de 25 à 30 %), mais l’image globale représente assez bien la proportion des

ZH en « bon » ou « mauvais » état probable pour l'IBGN, avec un décalage de l'ordre de 4 % dont il faut tenir compte.

Enfin, ces modèles permettent de visualiser les grandes tendances régionales et de faire ressortir les principales « forces motrices » génératrices d'impact, mais ils n'ont en contrepartie qu'un faible pouvoir explicatif et aucune capacité à mettre en évidence les processus d'impact. Le regroupement nécessaire des variables d'occupation du sol en grandes catégories ne permet pas de faire ressortir des facteurs gérables. Cette approche est relativement novatrice en bioindication. Des modèles comparables, utilisant des régressions logistiques pour prédire un « bon état écologique » au sens de la DCE à partir de l'occupation du sol, ont été développés en Irlande par Donohue *et al.* (2006).

Comparaison de deux hypothèses de limite du « bon état IBGN »

L'enseignement majeur de ces résultats est que, pour l'ensemble du territoire métropolitain, et avec l'hypothèse « haute » pour la limite BE, le pourcentage de tronçons de cours d'eau qui atteignent actuellement le « bon état IBGN » serait de l'ordre de 42 à 38 % selon les prédictions (corrigées ou non) des modèles d'extrapolation, contre 49 % observés dans les réseaux de suivi (tableau 3, page 10). Toutefois, si l'on accepte l'idée que les modèles fournissent une image moins biaisée que les réseaux actuels, c'est la valeur de 42 % fournie par les modèles corrigés qu'il faut retenir. Avec l'hypothèse « basse », qui correspond seulement à une baisse de 6 % pour l'EQR-IBGN, le pourcentage de bon état atteindrait 60 %. La principale conclusion à en tirer, c'est qu'environ 18 % des zones hydrographiques se trouvent en « situation limite », à un point IBGN seulement au-dessus ou au-dessous de la limite du BE, selon l'hypothèse choisie.

Par comparaison, l'Irlande, dont la limite de « bon état » pour les invertébrés apparaît plus exigeante que celle de la France, aurait actuellement 70 % du linéaire de ses principaux cours d'eaux en « bon état » (Donohue *et al.*, 2006), dans un contexte de pression anthropique évidemment plus favorable.

Pressions dominantes et tendances régionales

Les pressions liées à l'urbanisation apparaissent toujours fortement discriminantes pour l'atteinte

du « bon état IBGN » ; ces pressions recouvrent à la fois les pollutions domestiques et industrielles, mais aussi les altérations physiques, en particulier la chenalisation des rivières pour lutter contre les inondations des zones urbanisées. D'après les modèles, la probabilité de « bon état IBGN » dans l'hypothèse haute est quasi nulle au-dessus d'un seuil de 2 % de territoires artificialisés dans les bassins. Ce seuil apparaît beaucoup plus bas que ceux trouvés dans la littérature, mais avec d'autres définitions du bon état écologique et des surfaces urbanisées (Allan, 2004 ; Garcia Bautista *et al.*, sous presse). En revanche, avec des données comparables aux nôtres, Donohue *et al.* (2006) trouvent pour l'Irlande un seuil à 0,03 % qui apparaît peu réaliste en France.

Les impacts liés aux terres cultivées semblent provoquer des dégradations moins accentuées, mais spatialement très étendues ; ces impacts proviennent à la fois d'apports polluants (nutriments, pesticides), mais aussi d'altérations physiques (colmatage par l'érosion des sols, et chenalisation). Les trois quart des zones hydrographiques prédites en « bon état IBGN » dans l'hypothèse haute ont moins de 30 % de territoires agricoles dits « à fort impact » dans leur bassin, ce qui confirme, s'il en était besoin, l'existence d'un impact sur les cours d'eau lié à ce type d'agriculture. En revanche, cette probabilité de « bon état » ne devient nulle qu'au-dessus de 80 % de terres arables dans les bassins, ce qui veut dire que l'on peut trouver des cours d'eau en bon état même dans des bassins largement cultivés. Ces seuils sont du même ordre que ceux relevés dans la littérature (Allan, 2004 ; Garcia Bautista *et al.*, sous presse), bien que là encore, très supérieurs à ceux proposés par Donohue *et al.* (2006) pour l'Irlande. Mais le fort écart entre ces deux seuils montre surtout que les impacts liés à l'agriculture placent les cours d'eau en limite du bon état écologique sur l'IBGN. Ceci est confirmé par la répartition des zones hydrographiques prédites en « situation limite », qui se situent majoritairement dans les hydro-écorégions de plaine à forte activité agricole.

L'image globale fournie par ces modèles révèle une situation relativement équilibrée, avec *grosso modo* 40 % de zones hydrographiques en « bon état IBGN », 40 % en « mauvais état IBGN » et 20 % en « situation limite ». Si les zones de montagne (Massif Central, Vosges, Jura, Pré-Alpes du sud, Corse...) apparaissent globalement préservées, les régions agricoles de plaine sont en

« situation limite » ou déjà impactées, tandis que les zones urbanisées sont généralement prédites en « mauvais état ».

Conclusion

Ces résultats, diffusés préalablement à la Direction de l'Eau du MEDD, ont contribué à la réflexion pour le choix d'une limite de « bon état écologique » pour l'IBGN. C'est l'hypothèse « haute » qui a été adoptée comme base de classification provisoire sur l'ensemble du territoire métropolitain (MEDD, circulaire 2005). L'une des raisons de ce choix « ambitieux et réaliste » est qu'il apparaît possible, grâce à des mesures de faible ampleur mais généralisées, de regagner un ou deux points IBGN sur une grande partie du réseau dans un délai de 10 ans.

Ce choix se trouve maintenant conforté par la comparaison, réalisée au cours d'un processus d'inter-étalonnage requis par la DCE, des limites de « bon état écologique » proposées pour les invertébrés des rivières par l'ensemble des États membres. En effet, la limite adoptée par la France se trouve très proche de la moyenne des valeurs proposées et aucun ajustement ne devrait s'avérer nécessaire pour la limite du « bon état ».

En fonction de ce choix, il n'en reste pas moins qu'un effort significatif devra être réalisé pour

atteindre les objectifs de la DCE. La probabilité actuelle de ne pas atteindre le « bon état IBGN » est toujours fortement liée à l'urbanisation, vraisemblablement responsable des situations les plus dégradées. L'agriculture dite « à fort impact » (terres labourées et cultures permanentes) provoque sur les milieux aquatiques des impacts certes généralisés, mais dont l'intensité reste globalement limitée. La plupart des zones de grande culture se trouvent en limite du « bon état » sur l'IBGN, mais cela signifie en contrepartie que l'objectif de la DCE ne semble pas irréaliste, même dans les régions très agricoles, puisqu'une amélioration même légère permettrait d'atteindre l'objectif dans ces régions, du moins en ce qui concerne les peuplements d'invertébrés benthiques.

L'agriculture et l'urbanisation ont en commun de générer des rejets polluants – diffus ou ponctuels – mais aussi d'exercer des pressions physiques importantes – chenalisation, altération des berges – sur les cours d'eau. L'amélioration de l'état écologique passera donc nécessairement par des actions visant à réduire les rejets polluants à la source, au niveau des bassins, mais aussi par une gestion plus écologique des corridors rivulaires. Il reste aux scientifiques à développer des modèles qui permettront d'orienter au mieux les actions prioritaires pour atteindre ces objectifs ambitieux mais, apparemment, pas hors de portée. □

Remerciements

Ce travail a bénéficié du soutien financier du ministère de l'Écologie et du Développement durable (MEDD – Direction de l'Eau), et du projet Européen REBECCA (STREP6, Contrat SSPI-CT-2003-502158). Il s'inscrit dans le programme de la Zone atelier Bassin du Rhône (ZABR). Nous remercions vivement tous les hydrobiologistes des DIREN, qui ont fourni les données IBGN sur lesquelles repose cette étude. Merci également à Béatrice Dutang du GIE VALOREZ pour son aide à la préparation du manuscrit.

Résumé

La mise en œuvre de la Directive-cadre européenne sur l'eau implique le développement de modèles pressions/impacts permettant d'identifier les actions nécessaires à la restauration. Nous avons développé des modèles « large échelle », reliant l'indice invertébré aquatique (IBGN) à l'occupation du sol du bassin versant. La méthode des arbres de décision produit des classifications permettant l'extrapolation spatiale des résultats. Des modèles d'extrapolation spatiale validés ont été employés pour représenter à l'échelle nationale l'état écologique probable des cours d'eau, basé sur l'IBGN. L'urbanisation du sol est apparue comme premier facteur négatif d'impact suivi de l'agriculture (terres arables) ; les espaces naturels sont apparus comme facteur positif pour l'état écologique. Nous avons également utilisé ce modèle pour tester deux hypothèses pour la limite de bon état écologique et pour simuler les résultats à l'échelle du territoire national. Beaucoup de secteurs soumis à une pression agricole élevée (plaines et région méditerranéenne) sont classés en limite de « bon état » suivant l'hypothèse prise en considération.

Abstract

For the implementation of the European Water Framework Directive, there is a need of pressures/impacts models allowing to set the priority actions for the restoration. We developed "large scale" pressures/impacts models, linking aquatic invertebrate index (IBGN) to watershed land cover. We used decision trees method to generate classifications allowing spatial extrapolation. Validated spatial extrapolation models were used to represent on a national scale the probable current ecological status of rivers based on IBGN. Urban land cover appeared as the first negative impact factor followed by agriculture (arable land); natural areas appeared as a positive factor on ecological status. We used this model to test two hypothesis for the "good status" boundary, and to simulate the results at the national scale. Many areas under high agricultural pressure (lowlands and Mediterranean region) are in borderline situation according to the "good status" hypothesis.

Bibliographie

- AFNOR, 1992, *Essais des eaux. Détermination de l'indice biologique global normalisé (IBGN)*, NF T 90-350, ISSN 0335-3931.
- AGENCES DE L'EAU, 1999, *Système d'évaluation de la qualité des cours d'eau SEQ-EAU*, Les études des agences de l'eau, n° 64, 59 p.
- ALLAN, J.-D., 2004, Landscapes and riverscapes: the influence of land use on stream ecosystems, *Annual Review of Ecology, Evolution and Systematics*, vol. 35, p. 257-284.
- BREIMAN, L., FRIEDMAN, J.-H., OLSHEN, R.-A., STONE, C.-J., 1984, *Classification and regression trees*, Wadsworth International Group, Belmont, California, USA.
- BUFFAGNI, A., ERBA, S., BIRK, S., CAZZOLA, M., FELD, C., OFENBÖCK, T., MURRAY-BLIGH, J., FURSE, M.-T., CLARKE, R., HERING, D., SOSZKA, H., VAN DE BUND, W., 2005, *Towards European Inter-calibration for the Water Framework Directive: Procedures and examples for different river types from the E.C. project. STAR, 11th STAR deliverable*, STAR Contract N° EVK1-CT 2001-00089, IRSA, Roma.
- CHANDESRIS, A., WASSON, J.-G., PELLA, H., SAUQUET, E., MENGIN, N., 2006, *Typologie des cours d'eau de France métropolitaine. Appui scientifique à la mise en œuvre de la Directive Cadre Européenne sur l'Eau*, Cemagref, ministère de l'Écologie et du Développement durable, Lyon, 62 p.
- DONOHUE, I., MC GARRIGLE, M.-L., MILLS, P., 2006, Linking catchment characteristics and water chemistry with the ecological status of Irish rivers, *Water Research*, vol. 40, n° 1, p. 91-98.

GARCIA BAUTISTA, A., et al., sous presse, Combined pressures and geographical context, in REBECCA WP4 Rivers. Deliverable 6: Report on existing methods and relationships linking pressures, chemistry and biology in rivers, ANDERSEN J.-M., DUNBAR, M., FRIBERG, N., European Commission.

HANLEY, J.-A., MC NEIL, J., 1982, The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, vol. 146, n° 1, p. 29-36.

MEDD, 2005, *Directive-cadre européenne sur l'eau. Définition du « bon état » des eaux, constitution des nouveaux référentiels et des modalités d'évaluation de l'état des eaux douces de surface (cours d'eau et plans d'eau). Valeurs-seuils provisoires du « bon état » (période transitoire 2005/2007: programmes de mesures, choix des objectifs environnementaux.)*, Circulaire DCE 2005/12, ministère de l'Écologie et du Développement durable, Direction de l'Eau, Paris, 17 p.

LOUDIN, L.-C., MAUPAS, D., 2003, *Système d'évaluation de la qualité de l'eau des cours d'eau*, Rapport de présentation SEQ-Eau (version 2), ministère de l'Écologie et du Développement durable, Agences de l'eau.

PELLA, H., CHANDESRIS, A., WASSON, J.-G., 2004, Constitution d'un système d'information à référence spatiale dans le contexte de la Directive Cadre européenne sur l'eau, *Ingénieries-EAT*, n° 40, p. 11-20.

ROCHE, P.-A., BILLEN, G., BRAVARD, J.-P., DÉCAMPS, H., PENNEQUIN, D., VINDIMIAN, E., WASSON, J.-G., 2005, Les enjeux de recherche liés à la directive-cadre européenne sur l'eau, *Comptes Rendus Geoscience*, vol. 337, p. 243-267.

TENENHAUS, M., 1998, *La regression PLS, théorie et pratique*, Technip, Paris, 254 p.

VERNEAUX, J., GALMICHE, P., JANIER, F., MONNOT, A., 1982, Une nouvelle méthode pratique d'évaluation de la qualité des eaux courantes. Un indice biologique de qualité générale (I.B.G.), *Annales Scientifiques de l'Université de Franche-Comté (Besançon)*, 4^e série, n° 3, p. 11-21.

WALLIN, M., WIEDERHOM, T., 2003, *Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters*, REFCOND Guidance, Final version, EU Common Implementation Strategy (CIS) for the Water Framework Directive, Working Group 2.31 "Reference conditions for inland surface waters (REFCOND)", 86 p.

WASSON, J.-G., 2002, Les questions de recherche posées par la Directive Cadre Européenne sur l'Eau : problématique pour les eaux de surface continentales, *Hydroécologie Appliquée*, vol. 13, n° 1-2, p. 1-19.

WASSON, J.-G., CHANDESRIS, A., PELLA, H., BLANC, L., 2002, *Les hydro-écorégions de France métropolitaine. Approche régionale de la typologie des eaux courantes et éléments pour la définition des peuplements de référence d'invertébrés*, ministère de l'Écologie et du Développement durable, Cemagref BEA/LHQ, Lyon, 190 p.

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-EAT*, vol. 40, p. 3-10.

WASSON, J.-G., CHANDESRIS, A., PELLA, H., BLANC, L., VILLENEUVE, B., MENGIN, N., 2004, *Détermination des valeurs de référence de l'IBGN et propositions de valeurs limites du « bon état »*, rapport, Cemagref BEA/LHQ, Valorez, ZABR, Lyon, 69 p.

WASSON, J.-G., VILLENEUVE, B., MENGIN, N., PELLA, H., CHANDESRIS, A., 2005, *Modèles pressions/impacts : approche méthodologique, modèles d'extrapolation spatiale et modèles de diagnostic de l'état écologique basés sur les invertébrés en rivière (IBGN)*, rapport, Cemagref BEA/LHQ, Lyon, 61 p.